Lineage-affiliated specification events and molecular characterization during early stages of haematopoietic development

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

Vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Llucia Inès Albertí Servera

aus Palma (Mallorca), Spain

Basel, 2017

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Antonius Rolink
Prof. Dr. Daniela Finke

Basel, den 15.11.2016

Prof. Dr. Jörg Schibler
Dekan

“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.”

Marie Skłodowska Curie

Table of Contents

1. Abbreviations

ACK	ammonium-chloride-potassium
AML	acute myeloid leukemia
Batf3	basic leucine zipper transcriptional factor ATF-like 3
BCR	B-cell receptor
Blk	B lymphocyte kinase
BM	bone marrow
bp	base pair
BSA	bovine serum albumin
ccr2	chemokine receptor 2
CD	cluster of differentiation antigen
cDC	conventional DC
Cebpa	CCAAT/enhancer-binding protein alpha
Ciita	class II transactivator
CLP	common lymphoid progenitor
CMP	common myeloid progenitor
cor	correlation
CPM	counts per million
Csf1r	colony-stimulating factor 1 receptor
Ctla	cytotoxic T-lymphocyte-associate antigen
Cts	cathepsin
DC	dendritic cell
DEG	differentially expressed genes
DL1	delta-like 1
DN	double-negative
DNA	deoxyribonucleic acid
Dntt	DNA nucleotidylexotransferase
DP	double positive
dsDNA	double strand DNA
Ebf1	early B-cell factor 1

EDTA	ethylenediaminetetraacetic acid
EP	erythrocyte progenitor
EPLM	early progenitor with lymphoid and myeloid potential
ERC	extrachromosomal rDNA circles
F	fraction
FC	fold change
FACS	fluorescence-activated cell sorting
FB	follicular B cell
FcR	fragment crystallizable receptor
FDR	false discovery rate
FGF4	fibroblast growth factor-4
Flt3	Fms-like tyrosine kinase 3
Flt3L	Flt3 ligand
*Flt3L*tg	human *Flt3L* transgenic
FPKM	fragments per kilobase of transcript per million mapped reads
GO	gene ontology
HSC	haematopoietic stem cell
Igα	immunoglobulin alfa
Igll1	immunoglobulin lambda-like polypeptide 1
IL	interleukin
Il7r	IL7 receptor
IL7Rα	IL7 receptor alfa chain
IRES	internal ribosomal entry sequence
Itgb7	integrin beta-7
Lat	linker for activation of T-cells
Lax	linker for activation of X cells
Lck	leukocyte C-terminal Src kinase
LD	limiting dilution
Lin	lineage
LMPP	lymphoid primed multipotent progenitor
LSK	Lin^- $Sca\text{-}1^+$ $c\text{-}Kit^+$
LT-HSC	long term HSC

Ly	lymphocyte antigen
March1	membrane-associated RING-CH protein I
M-CSF	macrophage colony-stimulating factor
MEP	megakaryocyte-erythroid progenitor
MHC	major histocompatibility complex
MHCII	MHC class II
Mpo	myeloperoxidase
MPP	multipotent progenitor
mRNA	messenger RNA
MZB	marginal zone B cell
NGS	next generation sequencing
NK	natural killer
NSG	NOD scid gamma
PAM	partioning around medoids
Pax5	paired box protein 5
PBMC	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
pDC	plasmacytoid DC
Prtn3	proteinaise 3
qRT-PCR	quantitative reverse transcription PCR
QS	quality score
Rag	recombination-activating gene
RIN	RNA integrity number
RNA	ribonucleic acid
RNA-seq	RNA sequencing
SCF	stem-cell factor
sd	standard deviation
SDF1	stromal cell-derived factor-1
SEM	standard error of the mean
SiglecH	sialic acid binding Ig-like lectin H

Sla2	src-like-adapter protein
Stat5	signal transducer and activator of transcription 5
ST-HSC	short term HSC
TCR	T-cell receptor
TdT	terminal deoxynucleotide transferase
TF	transcription factor
Th2	T helper cell type 2
Tlr	toll-like receptor
TN	triple negative
Trat1	T-cell receptor-associated transmembrane adapter 1
UMI	unique molecular identifiers
VCAM-1	vascular cell-adhesion molecule-1
vs	versus
WT	wild type
Zap70	70 kDa zeta-chain associated protein

2. Summary

The well-established "classical" model of haematopoiesis reflects a hierarchical decision-making process where early multipotent progenitors make an irrevocable decision to differentiate towards either the lymphoid or myeloid lineages through so-called Common Lymphoid and Common Myeloid progenitor intermediates respectively. However, the proposals for alternative differentiation pathways and the description of progenitor cells that contradict the lympho-myeloid dichotomy have prompted multiple revisions of the strict compartmentalized classical model. We have previously characterized a $B220^{+}$ $CD117^{int}$ $CD19^{-}$ $NK1.1^{-}$ uncommitted and multipotent haematopoietic progenitor with combined lymphoid and myeloid potential that we called Early Progenitor with Lymphoid and Myeloid potential (EPLM). The emergence of high throughput methods enabling the investigation of single-cell whole-transcriptome profiles generates data that enhances the active debate regarding the heterogeneity of apparently phenotypically homogenous progenitors having different multiple lineage potentials. This thesis provides a detailed analysis of EPLM heterogeneity by combining the alternative and complementary "top-down" and "bottom-up" experimental designs. Using the "top-down" approach based on the expression of the cell-surface markers Ly6D, SiglecH and CD11c, we could subdivide EPLM into four subpopulations with distinct lineage biases. As revealed by the subsequent functional experiments, the $Ly6D^{+}$ EPLM fraction was lymphoid restricted and contained most B-cell potential whereas the so-called triple negative (TN) EPLM expressing none of the above markers remained as a lympho-myeloid multipotent fraction and the potential precursor of the $Ly6D^{+}$ subset. Subsequently, single-cell RNA sequencing ("bottom-up" approach) of 152 $Ly6D^{+}$ and 213 TN single cells revealed that in fact TN are composed of a mixture of cells where the myeloid potential is mainly due to the contribution of the G3 TN subset whereas the lymphoid potential resides in the G1 TN clustered group of cells. This heterogeneity was masked in previous bulk molecular and functional experiments, thus demonstrating the power of single-cell RNA-sequencing technology to study heterogeneity in haematopoietic progenitors at an unprecedented resolution. Moreover, single-cell transcriptome profiles enabled

the detection in an unbiased manner of markers that better define cellular identity. Here we redefined the "top down" EPLM classification by identifying Terminal deoxynucleotide Transferase (TdT) as a potential marker with which to discriminate the lymphoid and myeloid potential of EPLM since, in addition to the previously identified lymphoid primed Ly6D$^+$ cells, TdT is also expressed in the G1 TN fraction, which turned to be molecularly indistinguishable from the G2 Ly6D$^+$ fraction. The use of other candidate markers such as Ebf1 and CD115 enabled us to prospectively isolate cells from different newly identified subgroups of EPLM and to confirm their genetic signatures with functional assays, thus supporting the increasing belief that the repertoire of genes expressed reflects the immediate lineage bias of that cell. Finally, within the Ly6D$^+$ cells, we found a B-cell priming gradient and propose that the G1 Ly6D$^+$ fraction is the direct precursor of the first B-cell committed stage, the CD117$^+$ CD19$^+$ Pro-B cell. Therefore, we favour the concept that haematopoiesis occurs through a process of graded commitment where molecular priming is initiated earlier than previously anticipated. Overall, this study provides a valuable model demonstrating that previously characterized, phenotypically homogeneous, multipotent progenitor cells are in fact composed of mixtures of cells with differently restricted differentiation capacities.

3. Introduction

3.1 Initiation of haematopoiesis

Haematopoiesis is a well-orchestrated and tightly regulated process resulting in the formation of all blood cell types. In adult mammals, haematopoiesis starts in the bone marrow (BM) with a rare cell type called Haematopoietic Stem Cell (HSC) that has both self-renewal and multipotent capacities. Self-renewal is the ability to generate cells that are identical to the mother HSCs without differentiation, thus allowing the maintenance of the stem cell pool size, while multipotency is the ability to differentiate into all functional blood cells.

Self-renewal and differentiation of HSC requires a special environment, termed the hematopoietic stem cell niche, which provides the protection and the necessary signals for their maintenance and development [1, 2]. Two main microenvironments are suggested to form the bone marrow niches: the endosteal and the vascular niche [3]. The endosteal niche, located in the interface between bone and BM [4], includes osteoblasts, derived from mesenchymal precursors, that produce a number of cell-signalling molecules that contribute to the maintenance and regulation of HSC such as Jagged (ligand of Notch receptor), thrombopoietin and angiopoietin, SCF, chemokines or Wnt [5]. The vascular niche, located at the fenestrated endothelium of bone marrow sinusoids [4, 6], promotes proliferation and differentiation, active cycling and generation of short-term HSCs by producing factors important for mobilization, homing and engraftment of HSC such as vascular cell-adhesion molecule-1 (VCAM-1), stromal cell-derived factor-1 (SDF1) or fibroblast growth factor-4 (FGF4) [7]. An important function of the stem cell niche is to regulate the balance between cellular self-renewal and differentiation [2]. One mechanism that ensures this balance is the control of asymmetric/symmetric cell division. As a result of asymmetric division, one daughter cell maintains the mother stem cell fate and the other daughter cell becomes more committed to terminal differentiation [8, 9]. Several reported mechanisms argue for an intrinsic capability of HSCs to divide asymmetrically, exemplified by the unequal molecular segregation during cell division. Some examples are: the asymmetric distribution of regulatory

cell surface molecules (e.g. notch [10]) or cellular processes [11] in mammalian neurogenesis; asymmetric distribution of cellular organelles (e.g. p-granules in C.elegans [12]); asymmetric distribution of transcription factors and co-factors (e.g. Prospero/Miranda in Drosophila [13]); and asymmetric distribution of DNA (e.g. extrachromosomal rDNA circles, ERCs, in yeast [14]). Alternative to the intrinsic model, there is a post-mitotic model arguing that the observed differences in the cell fates of sister cells are due to post-mitotic mechanisms in which initially identically specified daughter cells influence each other's cell fate (e.g. via Notch signalling pathway [15, 16]). Further investigations are required to resolve whether both models co-exist and to what extent each model contributes to different specification of daughter cells.

Multi-potentiality of HSCs was first demonstrated by Osawa et al. after showing that a single mouse $CD34^{low/-}$ HSC reconstitutes haematopoiesis long-term in a lethally irradiated mouse [17, 18]. Similar results were reported by Notta et al. with single human HSCs engrafting NGS mouse [19]. Since most mature blood cell types are short lived, HSCs are ultimately responsible for the continuous and life-long renewal of blood. Mature blood cells are produced at an impressive rate of more than one million cells per second in the adult human.

The description of HSCs, and their isolation by fluorescence activated cell sorting (FACS), relies on the presence or absence of a range of cell-surface molecules. HSCs do not express markers that are associated with the various haematopoietic cell lineages (Lin^-) including, for example, CD3 (T lymphocytes), B220 and CD19 (B lymphocytes), CD11b (monocytes-macrophages), Ly-6G (neutrophils) and TER-119 (erythoid cells), and they express the two molecules c-Kit, a tyrosine kinase receptor for the cytokine stem cell factor (SCF), and Sca-1 a phosphatidynositol-anchored membrane protein of unknown function [20]. Thus, HSCs are enriched in the Lin^-, $Sca\text{-}1^+$, $c\text{-}Kit^+$ fraction of bone marrow cells, termed LSK cells, that represent approximately 0.05% of nucleated adult mouse BM cells [21]. The LSK compartment contains cells that repopulate the entire haematopoietic system either long-term (LT-HSC) or short-term (ST-HSC). Three different FACS staining strategies describe the phenotype of the two LSK subsets. The original Weissman approach relies on Thy-1.1 and Flt3 (fms-like) tyrosine kinase, a type 3-

receptor kinase, whose expression was observed to correlate with a loss of the capacity of HSCs to self-renew [22, 23]. In this model LT-HSCs are LSK $Flt3^{-}$ $Thy1^{low}$ and, ST-HSCs are LSK $Flt3^{+}$ $Thy1^{low}$. Later, in 2005, Jacobsen combined LSK markers with CD34 and Flt3 [22], a commonly used strategy. Thus Jacobsen et al. described LT-HSCs as LSK $CD34^{-}$ $Flt3^{-}$ and ST-HSC as LSK $CD34^{+}$ $Flt3^{-}$. Recently, Morrison made use of additional signalling lymphocyte activation molecule markers (CD150, CD48, CD229, and CD244) to subdivide LSK cells. Here, LT-HSCs are defined as LSK $CD150^{+}$, $CD48^{-/low}$, $CD229^{-/low}$ and $CD244^{-}$ while ST-HSC differ in CD229 expression and are defined as $CD150^{+}$, $CD48^{-/low}$, $CD229^{+}$ and $CD244^{-}$ [23].

HSCs differentiate into multipotent progenitor cells (MPPs), which gradually lose their multipotentiality while undergoing decision-making, expansion, and differentiation via intermediate lineage precursors giving rise to the final compartments of functional cells. The long-standing classical model of haematopoiesis [24], together with other versions [25], reflects a hierarchical decision-making process where early multipotent progenitors make an irrevocable decision to differentiate towards either the lymphoid or the myeloid lineage [26].

This clear lympho-myeloid separation is supported by the identification of progenitors of each lineage. Weissman and colleagues reported the so-called common lymphoid progenitor (CLP) [27] which can give rise to B and T lymphocytes and natural killer (NK) cells, and the common myeloid progenitor (CMP) [28] which can differentiate into the rest of the immune and blood cells. Thus, this traditional model illustrates single routes of differentiation to each end-cell type. However, over the past 15 years, the above strict compartmentalization of hematopoietic cells has been challenged. This is mainly due to the description of progenitor cells that contradict the lympho-myeloid dichotomy.

3.2 Early Progenitor with Lymphoid and Myeloid developmental potential (EPLM)

The first indication that haematopoietic development has much more plasticity than previously anticipated came from studies performed in mice deficient

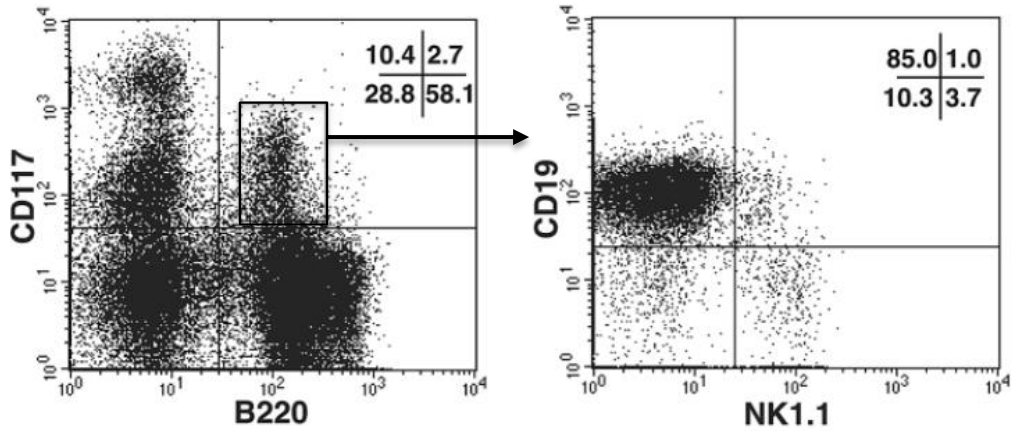


Figure 1. Gating strategy for the EPLM in the BM of adult C57BL/6: $B220^{+}$ $cKit^{int}$ $CD19^{-}$ $NK1.1^{-}$

for the transcription factor *Pax5* [29]. The $Pax5^{-/-}$ mouse presents an absolute block in B-cell development at the Pro-B cell stage. Remarkably, and in contrast to wild-type (WT) cells, *Pax-5* deficient Pro-B cells had multilineage developmental potential since they could develop into myeloid, NK, and T cells both *in vitro* and *in vivo* [30-32]. Moreover, even WT committed precursor B cells regain multipotentiality upon conditional inactivation of the *Pax5* gene [32]. This demonstrates that progenitor cells that have progressed some way along a given lineage retain the ability to give rise to other cell types until a rather advanced stage of differentiation. However, whether such a degree of developmental plasticity occurs under physiological conditions is an unresolved issue. Subsequently, our laboratory detected in wild-type mouse BM a cell with similar properties to the $Pax5^{-/-}$ Pro-B cells. This cell has the capacity to differentiate into lymphoid and myeloid cells and was therefore called EPLM (Early Progenitor with Lymphoid and Myeloid developmental potential). EPLM cells were identified as $B220^{+}$ $cKit^{int}$ $CD19^{-}$ and $NK1.1^{-}$ cells representing 0.2% of all nucleated BM cells [33] (**Figure 1**). In terms of phenotype, this progenitor is closely related to the CLP with the marked difference of B220 expression, EPLM being $B220^{+}$ whereas CLP are $B220^{-}$, and partially overlaps with the so-called Fraction A cells identified by Hardy and co-workers [34] (**Table1** contains a comparison of the nomenclatures used for early B-cell developmental stages). Limiting dilution analysis of EPLMs cultured together with stromal cells and addition of appropriate cytokines, enabled the quantification of *in vitro* B, T and, myeloid precursor frequencies. EPLMs showed strong B-cell developmental potential and strong-to-moderate differentiation potential for T cells and myeloid cells (mostly macrophages).

Therefore, this suggested that under physiological conditions the developmental fate of EPLM is mainly to become B cells. Reconstitution assays in order to assess the EPLM's *in vivo* developmental potentials revealed their ability to transiently reconstitute both B and T cell compartments in sublethaly irradiated *Rag2*-deficient mice.

Another cell that contradicts the classical lympho-myeloid lineage separation is the Lymphoid-Primed MultiPotent progenitor (LMPP), which has little potential for megakaryocyte or erythroid development while retaining other potentials [35]. Compared with EPLM, LMPP present slower kinetics in developing into lymphoid cells, hence suggesting an earlier and upstream location in haematopoietic development. Moreover, LMPP is the Flt3 positive high fraction of LSK and, in contrast with EPLM, is $B220^{-}$. Together with the identification of new progenitor cells, the finding of alternative routes to end-cell types [36] prompted the necessity of proposing new models for the architecture of haematopoiesis.

Developmental stages			
Osmond	Melchers and Rolink	Hardy	Status of Ig loci
Pro-B	Pre-pro B / EPLM	A	Germline
	Pro-B	B	D-J_H rearrangement
		C	V_H-DJ_H rearrangement
Pre-B	Large pre-B	C'	V_HDJ_H pairs with λ5-Vpre-B
			Pre-BCR surface expression
	Small pre-B	D	$V_κ$-$J_κ$ or $V_λ$-$J_λ$ rearrangement
Immature B	Immature B	E	Complete BCR (receptor editing can occur)

Table1. **B-cell developmental stages in the bone marrow**. Comparison of the nomenclatures used to identify developmental B cell subsets and how they relate to key VDJ recombination events (comprehensively reviewed in Osmond et al. [37], and Hardy et al. [38]). Table adapted from Hardy et al. [39].

3.3 Pair-wise model of haematopoiesis

In collaboration with others, a few years ago our laboratory presented a new model of haematopoiesis called the pair-wise model [40, 41]. In contrast to the traditional hierarchical trees, this model outlines haematopoietic development in a cyclical representation. Fundamentally, the principle of an HSC giving rise to all blood cell types is the same. However, the intermediate progenitor stages are better illustrated. In the pair-wise model, the set of developmental potentials that have

been shown experimentally are represented as a continuum by arcs (**Figure 2**). As development proceeds, the arcs become shorter indicating that developmental potentials are gradually restricted and ultimately cells become committed to a differentiated cell fate (as shown by the outermost arc). As its name indicates, this model depicts a series of pair-wise developmental relationships between the various haematopoietic lineages. This infers that only certain bi-potentialities are permissive. For example, a bi-potent T lymphocyte-megakaryocyte cell should not exist, and has not been described to date.

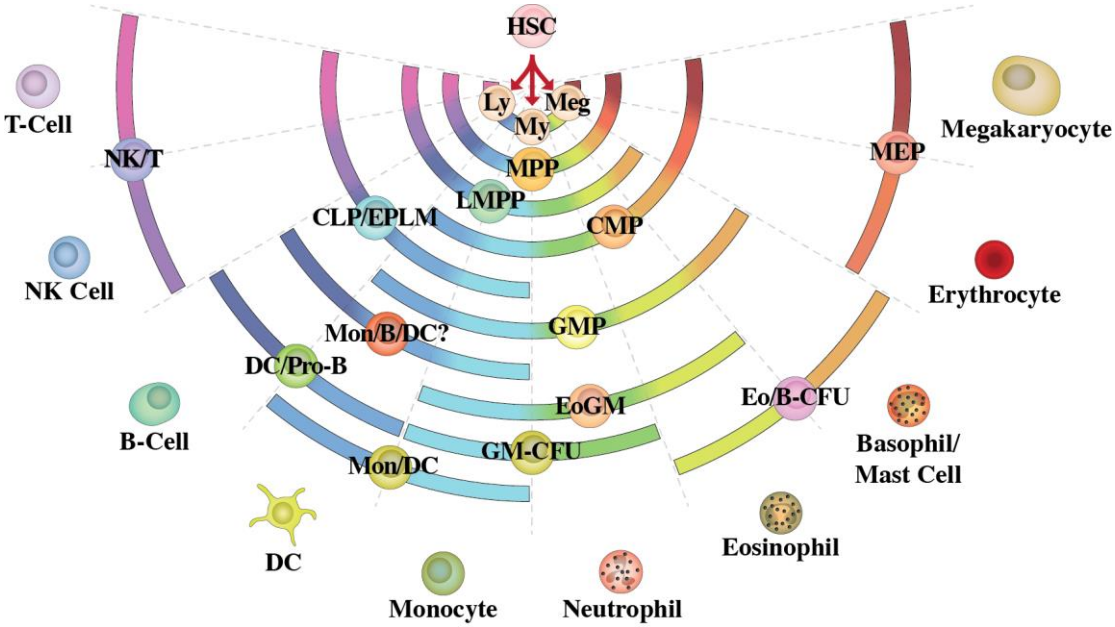


Figure 2. Pair-wise model of haematopoiesis. Illustrated is a fate choice continuum with an invariant series of pair-wise developmental relationships between haematopoietic cells' fates.

The pair-wise model has no lineage branching patterns and therefore, does not assume a single preferred route to a particular end-cell type. In other words, HSCs can reach a specific differentiated blood cell through more than one type of intermediate progenitor. For instance, dendritic cells (DC) can be derived from either CLP/EPLM or CMP, indicating that DC can have both lymphoid and myeloid origin [36]. Moreover, multiple routes have been demonstrated towards neutrophils and monocytes [42, 43]. Another advantage is that the pair-wise model maintains the possibility that progenitors can move in the opposite direction and regain multipotentiality -as is the case for the $Pax5^{-/-}$ Pro-B cells. Overall, the above-described model, which includes the recently identified progenitors and allows

alternative routes, leads to a viewpoint that haematopoiesis is a more versatile and less strictly compartmentalized process than previously thought.

In line with the description of progenitor cells that have multiple lineage options, there is an increasing debate regarding their heterogeneity. It is highly likely that multipotent progenitor cells presently viewed as a homogeneous population will be divided into subsets with lineage biases or even into already committed subsets. In turn, some end-cell types will be probably further divided into numerous subtypes. To what extent progenitor cell populations, and their mature progeny, are heterogeneous is very much dependent on the quantity of new and existing cell-surface markers that can be used to define new subpopulations. In this regard, together with the improvement of traditional flow cytometer-based technologies, allowing the simultaneous detection of an increasing number of markers, the explosion of high throughput technologies, such as RNA sequencing that enable the parallel screening of a large amount of markers in an unbiased manner, has been of a pivotal importance.

3.4 Single-cell RNA sequencing: the method that dissects heterogeneity

In the last couple of years, RNA sequencing at the single-cell level has emerged as an important technology for interrogating the degree of heterogeneity in a population. This technology, which can be performed at a whole-transcriptome scale with unprecedented resolution, consists of extracting the RNA from individual cells followed by a sequencing process similar to that for bulk RNA sequencing. When RNA sequencing is performed at the population level, as has been the case for most studies, we obtain an average transcription profile of thousands or millions pooled cells. This, masks the true distribution of gene expression levels across individual cells, thus hampering the elucidation of whether a subset of cells is responsible for the expression of a given gene or even if it is the major contributor of a function in a seemingly homogeneous population [44]. Therefore, it is crucial to go to the single-cell level where there is no averaging effect.

3.4.1 Technological advances

The first mRNA sequencing whole-transcriptome analysis of a single cell was reported in 2009 by Tang and co-workers. Using a novel mRNA sequencing assay they detected expression of 5,270 genes in a single mouse blastomere [45], thus greatly overcoming the limited throughput of a couple of hundred genes that can be analysed by single-cell qRT-PCR. Moreover, in contrast to qPCR, mRNA sequencing based technology provides a comprehensive view of the transcriptome profile in one cell in an unbiased manner or, in other words, without requiring prior knowledge and selection of genes of interest. However, the method reported relies heavily on manual manipulation of cells and reagents and can only process few cells per hour. The limitation of manual handling, either during the capturing or during the later processing steps, also occurs in other single-cell RNA sequencing technologies such as laser-capture microdissection or single-cell sorting into a multi-well plate. Therefore, researchers rapidly attempted the development of automated methods.

Automation of the single-cell RNA sequencing process became a reality with the emergence of chip-based microfluidics devices. In 2013, Fluidigm launched the "C1™ Single-Cell Auto Prep System" helping expand access to the single-cell RNA sequencing technology [46, 47]. The C1 system provides a completely automated workflow that integrates both the capture of single cells in a microfluidic chip and the later processing steps including washing, lysis, reverse transcription and, PCR pre-amplification. In addition, the major strength of microfluidics is the manipulation of minute volumes, working in a nanoliter reaction scale. As a result, albeit allowing the processing of only a few tens of cells per experiment, this system presents higher reproducibility, more cost-effective reactions and increased sensitivity compared with classical methods [48]. In parallel with its commercialization, the journal "Nature Methods" declared single-cell sequencing as the method of the year 2013. As a consequence, the past few years have seen a burst of papers from laboratories around the world that have developed expertise in single-cell approaches. These papers include new single-cell technologies and provide evidence for its valuable application in many areas.

Regarding the development of new technologies, the state-of-the-art has shifted towards scaling up the number of cells that can be analysed in parallel and in an affordable manner in terms of time and costs. Recently, Klein et al. and Macosko et al. have developed two advanced droplet-based microfluidics methods termed Drop-seq [49] and inDrop sequencing [50] respectively. Both methods, published in the same issue of the journal Cell, massively increase the throughput to several thousand of cells while reducing the cost of sequencing. Moreover, despite there being some differences between the experimental protocols, both publications rely on the same fundamental principle. The cells are separated into nanoliter-sized aqueous droplets in carrier oil as reaction chamber. Each reaction chamber contains microparticles coated with unique cell barcodes and Unique Molecular Identifiers (UMIs). Although droplet microfluidics is not a new technology, this efficient barcoding strategy makes it an attractive method and, altogether, has big potential for commercialization.

To date, apart from droplet-based microfluidics, a variety of other microfluidics techniques for single-cell trapping exist. Some examples are hydrodynamic trap [51], magnetic trap [52, 53], acoustic trap [54], dielectrophoretic trap [55], and optical trap [56]. All this together exemplifies increasing interest for single-cell handling technologies and creates anticipation for exciting biological advances for the coming years.

3.4.2 Applications

Single-cell transcriptomics offers a wide range of applications that have already enabled some advances in a broad variety of fields apart from immunology. Examples include: oncology, with the dissection of tumour cell heterogeneity [57], the identification of subgroups associated with anti-cancer drug resistance [58], and the analyses of rare circulating tumour cells [59] or cancer stem cells; embryology, with new insights into the gene regulatory networks controlling early stages of embryonic development [45, 60-62]; neurobiology, with the unravelling of the complex cellular composition of the mouse hippocampus while uncovering novel markers and cell types [47]; or microbiology, with the opportunity of sequencing

uncultivated microbial eukaryotes at the single-cell level and identifying new species, unknown microbial functions and revised evolutionary relationships between eukaryotes, bacteria and archea [63, 64].

In immunology, single-cell molecular profiling is providing valuable information on the differentiation and function of the immune cells both at the steady state and during immune responses. Moreover, it is opening up a new avenue for finding markers that better describe cell phenotypes in an unbiased manner. During the forthcoming years, it is predicted that the identity of many immune cells will be redefined while possibly new cell types and subtypes will emerge. Indeed, several studies have already pointed in this direction. Mahata et al. identified a Th2 subpopulation contributing to the maintenance of T-cell homeostasis. It is of interest that the Th2 subset, characterized by the expression of a specific enzyme (Cyp11a1), could be isolated by using surface molecules whose expression correlates with enzyme expression, thus allowing for the first time a broad *ex vivo* functional validation of a new cell type discovered by single-cell RNA sequencing [65]. Similarly, Shalek et al. found a very small subset of cells among mouse bone-marrow-derived dendritic cells. Those cells, termed precocious, are the first ones that produce and secrete a wave of interferon during antiviral responses [66]. Another example with novel subsets is a recently published paper providing a detailed study about the inter-cellular transcriptomic variation within the traditionally classified classical, intermediate and non-classical monocytes [67]. In the publication, Gren et al. reveal that each group of monocytes contains further subgroups with distinct genetic signatures according to their activation status and differentiation. This indicates the ability of single-cell transcriptomics to discover cell heterogeneity within defined cell populations.

However, as a proof of concept, massive parallel single-cell RNA sequencing has also been applied to more complex and heterogeneous systems such as the spleen [68] or peripheral blood mononuclear cells (PBMCs) [53]. Apart from the identification of previously unknown hidden subpopulations, this approach could reconstitute the global cell heterogeneity within splenic tissue and PBMCs, thus envisaging the opportunity to redefine our knowledge on lineage hierarchy. In fact, a number of studies have amplified the current hematopoietic classifications and

differentiation trajectories through the detection of transient or intermediate cell states [69-71]. For instance, Paul et al. identified seven subgroups of c-Kit^{+} Sca1^{-} Lin^{-} sorted cells representing a gradient of erythrocyte transcription, from expression of early progenitor genes towards the up-regulation of functional genes [69]. It is of note that, unlike current models, none of the seven subsets co-expressed megakaryocyte genes. Therefore, they suggested that the standard gating and sorting for megakaryocyte-erythrocyte progenitors (MEP) might be termed erythrocyte progenitors (EP) instead. This exemplifies the necessity of revising the identity of the different immune cells as well as their subtypes, lineages and composition within tissues.

In conclusion, the studies reported to date have convincingly demonstrated that single-cell RNA sequencing is a powerful approach that can be used in many applications such as exploring cellular heterogeneity in a population, studying differentiation dynamics, redefining cell identity, or identifying rare cell types among others. Excitingly, it is expected that application of this technology will continue to expand as the techniques and analysis tools evolve.

3.5 The role of IL-7 and Flt3L cytokines in lymphocyte development

Cytokines are the pivotal external factors that deliver environmental signals to control haematopoietic cell development. Their mode of action can be viewed as either instructive or permissive [72]. In the instructive model, cytokines act directly on progenitor cells to drive them towards a specific lineage by actively inducing the initiation of a lineage-specific gene program. In contrast, in the permissive model commitment occurs in a cell-autonomous and stochastic manner and cytokines act to selectively allow cells committed to a particular lineage to survive and/or proliferate [73, 74]. The precise instructive versus permissive role of cytokines in haematopoiesis remains controversial [75-77]. Two important cytokines for T- and B–lymphocyte development are IL-7 and Flt3L since their receptors are co-expressed in a narrow window immediately before commitment and mutant mice show impaired lymphoid development [78-81].

IL-7 was originally identified as a growth factor for B-cell progenitors [82]. Later on, IL-7 activity on thymocytes and T-cell survival was also reported [83]. Subsequently, a specific receptor for IL-7, expressed in both B- and T-cell progenitors, was identified [84]. Mice lacking either the cytokine [81] or its receptor (IL-7Rα) [80] have a leaky arrest of T-cell development at the DN2 stage [80, 81] and absence of γδT cells [85], whereas a dramatic defect in B-cell generation with a block at the Pro-B cells stage occurs. Reports with over-expression of the pro-survival gene *Bcl2* demonstrated that whereas the T-cell defect is rescued, thus suggesting a permissive role of IL-7 in early T-cell development [86, 87], B-cell development is not re-established [88, 89], arguing for an instructive role of this cytokine in B-cell commitment at least in mice. Interestingly, humans with defects in IL-7 signalling display a normal B-cell development [90, 91]. However, an *in vitro* system has provided clear evidence that IL-7R signalling through Stat5 alone is not sufficient to induce transcription of *Ebf1* and *Pax5* [92], indispensable for B-cell commitment, therefore arguing for an important but redundant action of IL-7 during commitment to the B-cell lineage [93, 94]. Therefore, while the importance of IL-7 as a survival and growth factor for committed B-cell progenitors is well established, its instructive role through *Ebf1* and *Pax5* up-regulation remains unclear.

Flt3 ligand (Flt3L) is one of the essential cytokines for cell survival and proliferation during early haematopoiesis [95, 96]. Its function has gained much attention since mutations in Flt3L signalling are commonly found in Acute Myeloid Leukaemia (AML) [97]. Using flow cytometry, its receptor (Flt3) has been detected as early as the non-self-renewing MPP stage of development [98]. Thereafter, several downstream myeloid and lymphoid progenitors continue to express Flt3 with the exception of megakaryocyte-erythrocyte progenitors [99]. Flt3 expression is extinguished upon lineage commitment and, among differentiated cells, only dendritic cells retain Flt3 on their surface. For instance, during B-cell commitment Pax5 expression directly antagonizes that of Flt3 [100]. In the thymus, Flt3 expression is used to identify the earliest and uncommitted thymocyte subpopulation with remaining B-cell developmental potential [101-103]. Mice with targeted gene disruption of *Flt3* [78] or its ligand [79] do not exhibit a complete loss of any haematopoietic population but reduced numbers of B, DC and NK cells [78,

79]. Moreover, upon transplantation, BM progenitors from *Flt3*$^{-/-}$ or *Flt3L*$^{-/-}$ mice reconstitute the B-cell compartment poorly [78, 79] and Flt3L was found to be essential for maintaining normal numbers of uncommitted B-cell progenitors [104]. Therefore, these results suggest a survival and proliferative action of Flt3L during B-cell development, rather than instructive. It is probable that Flt3L signalling exerts its function in concert with signalling from other cytokines such as SCF or IL-7 [92, 105]. This is manifested in cultures of lymphoid progenitor cells, such as ETP, CLP or EPLM, in the presence of IL-7, where Flt3L provides and additive anti-apoptotic effect while stimulating proliferation [106, 107].

3.6 *Flt3L*tg mouse

Until recently, both bulk and single-cell RNA sequencing required a considerable amount of starting material. For bulk experiments, it was about 500 nanograms of RNA per sample, while for the single-cell capturing step using the C1 platform the starting density recommended in the "single-cell preparation guide" from fluidigm is 166-255 cells/µL and 5-20 µL are loaded into the chip. To overcome this limitation, because the EPLM is a rare population, we made use of a mouse model, the *Flt3* ligand transgenic (*Flt3L*tg) mouse, to isolate EPLM and other progenitor populations in greater numbers.

Initially, Ceredig et al. showed that apart from DC, increased in vivo Flt3L availability after several injections led to transient expansion of Flt3^{+} progenitors with either myeloid or lymphoid developmental potential or both [108]. Later on, our laboratory generated a transgenic mouse with sustained over-expression of human Flt3L, thus providing for the first time a detailed *in vivo* analysis of the effect of this cytokine on different haematopoietic lineages [109]. Examination of *Flt3L*tg mice confirmed the importance of Flt3L in DC development. Moreover, the transgenic mice presented a tremendous expansion of almost all haematopoietic progenitors in the BM with the exception of MEP. The decrease in MEP occurred in a rapid fashion when WT mice were injected with recombinant Flt3L, therefore suggesting for the first time an instructive role for the Flt3L cytokine. Tsapogas and co-workers proposed that an increased level of Flt3L guides haematopoietic

development towards the lymphoid-myeloid fate at the expense of the megakaryocyte-erythrocyte fate [41, 109]. This provides an explanation for the decreased platelet counts and the consequent development of anaemia in *Flt3L*tg mice.

The EPLM compartment is a good example of the tremendous expansion of progenitors in the BM of *Flt3L*tg mice. In fact, the number of EPLM cells was increased by 14-fold when compared to WT mice (**Figure 3**). Therefore, the *Flt3L*tg mouse is an excellent tool for the isolation of large numbers of EPLM and other progenitor cells for further *in vitro* and *in vivo* experiments, transcriptome profiling or molecular biology analyses.

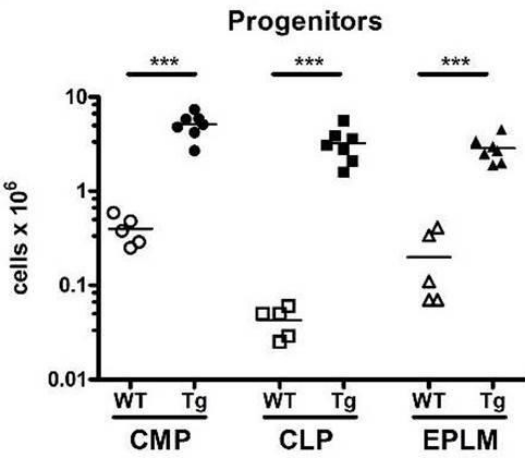


Figure 3. Total numbers of CMP, CLP and EPLM in the BM of WT (white symbols) and *Flt3L*tg (black symbols) mice (5-7 mice per group). ***P<0.0001.

4. Aim of the project

In this thesis we sought to characterize a previously identified uncommitted and multipotent progenitor cell called EPLM with the aim to:

- Dissect EPLM heterogeneity by combining the use of cell surface markers, such as Ly6D, SiglecH and CD11c, the so-called “top down” experimental strategy, with single-cell RNA sequencing, “bottom up” approach.
- Unravel whether EPLM phenotypic or molecular heterogeneity causes different sets of potentials in the identified subpopulations by *in vitro* limiting dilution assays and *in vivo* reconstitution experiments.
- Study the precursor-product relationship among the subpopulations in bulk culture with cytokines.
- Identify markers that better define cellular identity of the EPLM subpopulations in an unbiased manner, by single-cell RNA sequencing.
- Utilize the identification of EPLM subpopulations in order to study the role of cytokines (Flt3L and IL-7) in B-cell commitment (von Muenchow et al. **Appendix paper 1**).

5. Materials and methods

5.1 Mice

C57BL/6 (B6), B6 *Rag2*-deficient [110], B6 FLT3L transgenic (*Flt3L*tg, [109]) and *Ebf1*$^{ihCd2/+}$-*Flt3L*tg mice used herein were 6 to 11 weeks old and matched by age and sex for each experiment. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. All animal experiments were carried out according to institutional guidelines (authorization numbers 1886 and 1888 from Kantonales Veterinäramt, Basel).

Ebf1$^{ihCd2/+}$ mice were provided by Prof. Meinrad Busslinger and co-workers [111]. These mice had been generated by inserting an internal ribosome entry sequence (IRES)-*hCd2* (ihCd2) reporter gene into the 3' untranslated region of the *Ebf1* gene. We next crossed the *Ebf1*$^{ihCd2/+}$ reporter to the *Flt3L*tg mice, thereby generating the *Ebf1*$^{ihCd2/+}$-*Flt3L*tg mice.

5.2 Flow cytometry and cell sorting

Bone marrow cell suspensions were obtained from femurs of the two hind legs of mice. Bones were flushed with a 2ml syringe filled with PBS containing 0.5% BSA and 5mM EDTA. Afterwards, single-cell suspensions were subjected to ACK treatment for erythrocyte depletion, stained with the appropriate combination of antibodies for 30 minutes at 4°C, and washed for subsequent flow cytometry or cell sorting. The following antibodies were used (from BD Pharmingen, 18 eBioscience, BioLegend, or produced in house) with names in brackets describing the corresponding clone: anti-B220 (RA3-6B2), anti-CD117 (c2B8), anti-CD19 (1D3), anti-NK1.1 (PK136), anti-SiglecH (551), anti-CD11c (HL3), anti-Ly6D (49-H4), anti-CD115 (AFS98), anti-hCD2 (RPA-2.10) conjugated with FITC, PE, PE/Cy7, APC, Bv421 or Biotin plus streptavidin-Bv650. For detection of TdT, cells were fixed and permeabilized after cell-surface staining using the Foxp3 Fix/Perm buffer kit (eBioscience), and subsequently stained with APC-conjugated anti-TdT (19-3) according to the supplier's protocol. Flow cytometry was performed using a BD LSR

Fortessa (BD Biosciences) and data were analyzed using FlowJo v9.8 Software (Treestar). For cell sorting, a FACS Aria IIu (BD Biosciences) was used (>98% purity).

5.3 *In vitro* limiting dilution assay

ST2 [112], OP9 [113] and OP9 stromal cells expressing the Notch ligand Delta-like 1 (OP9-DL1) [114] were plated in a 96-well flat-bottom plate one day prior to co-culture at $3x10^3$ cells per well. The following day, semi-confluent stromal cells were γ-irradiated with 3000 rad using a Cobalt source (Gammacell 40, Atomic Energy of Canada, Ltd) at 100 rad/min and co-cultured with graded numbers of sorted haematopoietic progenitors in 48 replicates (or as indicated). Cells were maintained as a monolayer in IMDM supplemented with $5x10^{-5}$M β-mercaptoethanol, 1mM glutamine, 0.03% w/v Primatone (Quest Naarden, The Netherlands), 100U/mL Penicillin, 100 μg/mL Streptomycin and 5% FBS (Amimed) at 37°C in a humidified atmosphere containing 10% CO_2 in the air. OP9 and OP9-DL1 co-cultures were additionally supplemented with 10% IL7-conditioned medium. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), wells were inspected using an inverted microscope. Wells containing colonies of more than 50 cells were scored as positive. For each experiment, the frequency of negative wells was plotted against the number of haematopoietic progenitors plated and the fraction of progenitor cells developing B-cell, T-cell or myeloid colonies was estimated considering plating efficiency that follows a Poisson distribution.

5.4 *In vivo* reconstitution assay

Recipient *Rag2*-deficient mice were γ-irradiated using a Cobalt source as previously described at a dose of 400 rad 4h prior to reconstitution. Indicated numbers of sorted haematopoietic progenitors from donor mice (B6, *Flt3L*tg or *Ebf1*$^{ihCd2/+}$-*Flt3L*tg as specified) were injected intravenously. After 3 weeks, spleen and thymus of recipient mice were separately analysed by flow cytometry.

5.5 Bulk cultures with cytokines

A number of $5x10^4$ sorted haematopoietic progenitors from *Flt3L*tg mice were cultured with 50ng/ml Flt3L prepared in-house and 10% IL-7 conditioned medium in a 24-well flat-bottom plate. Cells were maintained as previously indicated and from day 2 to day 6, one well containing cells from each population was analysed by flow cytometry for Ly6D and CD19 expression.

5.6 Statistical analysis

Statistical analysis was performed with GraphPad Prism v6.0f software. Two-tailed unpaired Student *t* tests were used for statistical comparisons. Data are presented as mean values ± SEM (n.s. not significant or $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$).

5.7 Bulk RNA sequencing

5.7.1 Sample acquisition

$Ly6D^+$ and TN EPLM subpopulations as well as $CD115^-$ TN and Pro-B cells were sorted from femurs of 2-pooled male *Flt3L*tg mice (6 to 8 weeks of age). After each sort, cells were centrifuged, resuspended in 0.5ml of TRIzol reagent and stored at -80°C for later total RNA extraction. From the $Ly6D^+$ and TN samples, 100µl containing ~$3 x10^4$ cells were previously used for the capture of single cells.

5.7.2 RNA extraction

Total RNA was extracted from *ex-vivo* sorted samples using TRIzol-based method [115, 116]. Briefly, $1x10^5$ to $3x10^5$ cells were lysed in 0.5ml of TRIzol reagent and 0.1ml of chloroform was added per 0.5ml TRI reagent. After incubation and centrifugation for phase separation, the aqueous phase containing the RNA was recovered and mixed with isopropanol in a 1:1 ratio for RNA precipitation. Following 15min incubation and centrifugation, the supernatant was discarded while the RNA pellet was first washed with 75% ethanol and subsequently resuspended with 20µl

of DEPC treated water. Concentration and 260/280 purity ratio was initially determined using NanoDrop 1000 Spectrophotometer (Witec AG). Selected RNA samples were stored at -80°C for later usage.

5.7.3 RNA quality

Either 500ng or 1µg of total RNA per sample was sent to the Genomics Facility at the D-BSSE (Basel) for quality control, library preparation and sequencing. Quality and level of degradation of the extracted RNA was assessed with RNA integrity number (RIN) assigned by the Agilent 2100 Bioanalyzer instrument using either the Nano or the Pico Agilent RNA 6000 kit (Agilent Technologies). Samples with a RIN value over 8 and presenting clean peaks were considered for further analysis. The RNA quantity was measured by the Infinite M1000 PRO - Tecan instrument using the Quant-iT RiboGreen RNA Assay Kit.

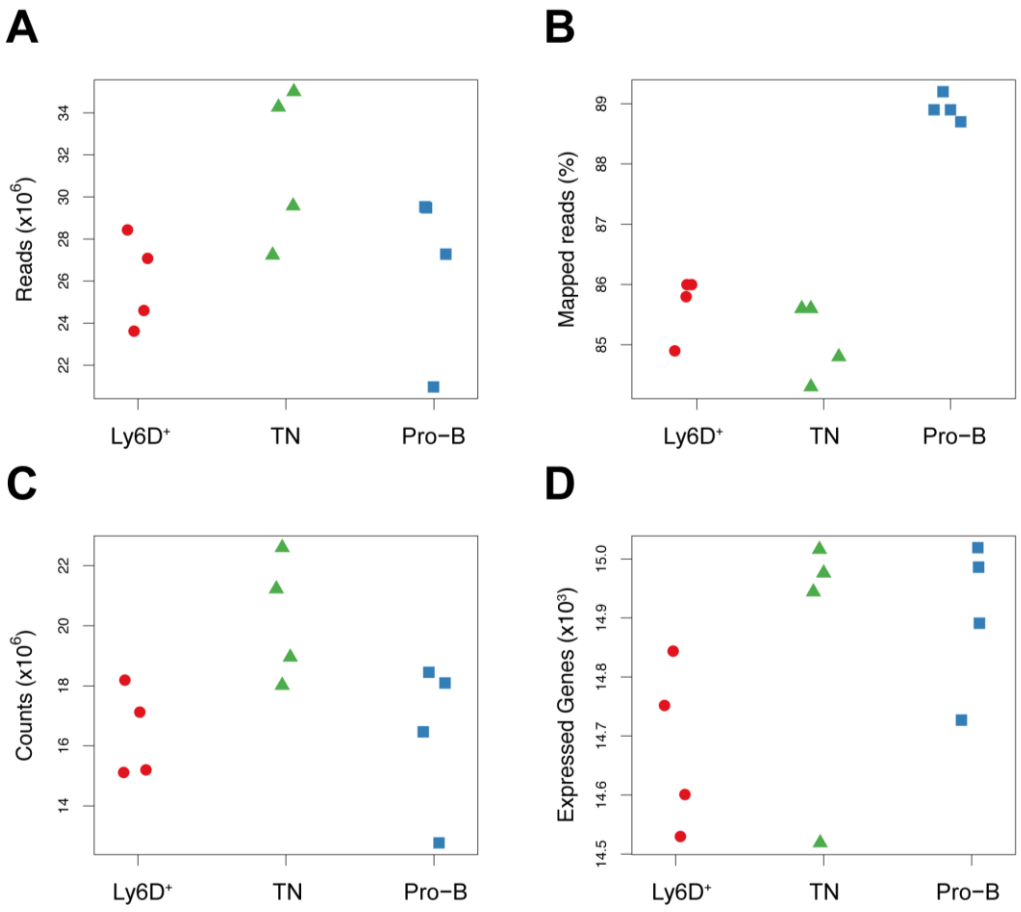


Figure M1. Quantification of raw sequenced data for the Ly6D$^+$, TN and Pro-B samples. (A) Number of sequenced reads per sample. **(B)** Percentage of reads that mapped to the mouse genome (mm9). **(C)** Number of counts per sample (library size) considering reads mapped to genes (exons only). **(D)** The total number of detected genes per sample (with at least 1 count).

5.7.4 Library preparation

For the preparation of sequencing libraries, the TruSeq Stranded mRNA LT Sample Preparation kit was used following the manufacturer's guide [117]. The polyA containing mRNA molecules were purified using poly-T oligo attached magnetic beads and subsequently fragmented using divalent cations under elevated temperatures. Afterwards, the RNA fragments were copied into first strand cDNA using reverse transcriptase and random primers. Strand-specificity information was achieved by replacing dTTP with dUTP during the second strand cDNA synthesis. To prevent self-ligation of the double-stranded cDNA, the 3' ends of the blunt fragments were adenylated followed by ligation of barcoded adapters suitable for Illumina-based sequencing. The product was subjected to 15 cycles of PCR amplification.

Size and purity of the library fragments was assessed by the Fragment Analyzer using the NGS Fragment 1-6000bp method (average fragment size 321bp, sd 20.36), while quantification was done with Quant-iT PicoGreen® dsDNA Assay Kit; TEcan instrument.

5.7.5 Sequencing

Indexed DNA libraries were pooled in equal volumes and loaded on one NextSeq 500 High Output flow cell (Illumina). Single-end sequencing was performed on the Illumina NextSeq™ 500 Sequencing System (D-BSSE, Basel) for 81 cycles yielding in 21 to 35 millions of reads, 81-mers, per sample **(Fig. M1 A).**

The Genomics Facility with the Illumina pipeline performed de-multiplexing and reads were transferred in the FastQ format via openBIS. A quality control of the sequenced data was performed using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, Version 0.11.3). All samples comprised high number of reads (> 21 millions) with median Quality Score (QS) of 35, a GC content distribution equivalent to the expected theoretical distribution (~52%), a sequence duplication level typical for RNA-seq samples, and no adapter content present (no need for trimming of reads). **Fig. M2** shows a representative example.

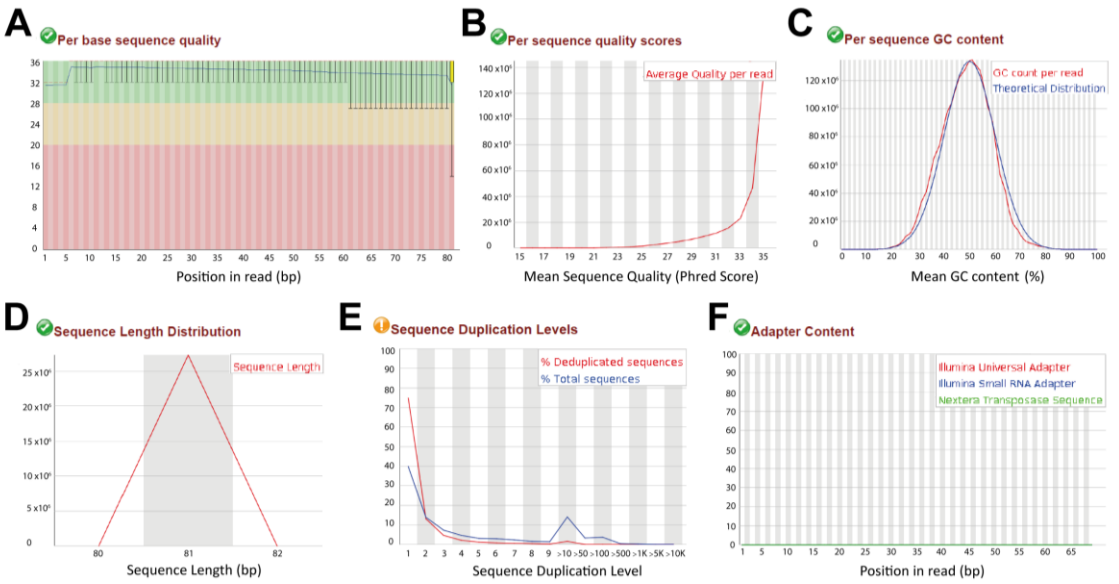


Figure M2. Example of quality control of raw sequence data (FASTQC). (A) Quality scores for individual positions within read sequence (over all reads). **(B)** Quality score distribution over all sequences. **(C)** GC content distribution over all sequences. Red: theoretical GC content (%); blue: observed GC content (%). **(D)** Distribution of sequence length over all sequences. **(E)** Relative number of sequences with different degrees of duplication. **(F)** Frequency of contamination by sequencing adapters. Replicate 2 of Ly6D group is taken as a representative example.

5.7.6 Pre-processing of sequencing data

All downstream analysis was performed using the open source software R accessed via RStudio server (R version 3.2.0). Sequencing reads were aligned to the mouse genome assembly, version mm9 (downloaded from UCSC http://genome.ucsc.edu), with SpliceMap [118, 119], included in the R/Bioconductor package QuasR, version 1.10.1. Splice-map was also capable of mapping reads that cover exon junctions. More than 80% of total reads were successfully mapped for each sample **(Fig. M1 B).** Subsequently, a count table with gene expression levels was generated using the qCount function from QuasR package and coordinates of RefSeq mRNA genes (http://genome.ucsc.edu, downloaded in December 2013). The expression level was defined as a number of reads that started within any annotated exon of a gene (exon-union model). Total counts per sample ranged from 13 to 22 millions **(Fig. M1 C)**, the so-called library size. Genes with no counts across all samples were filtered out from the analysis. For 17,290 genes at least 1 read was detected across all samples, corresponding to ~14,800 genes per sample **(Fig. M1 D)**.

Raw counts were normalized between samples with the TMM method (weighted trimmed mean of M-values [120]), expressed as counts per million mapped reads (CPM), and transformed to the log2-scale ($\log_2$CPM).

5.7.7 Data analysis

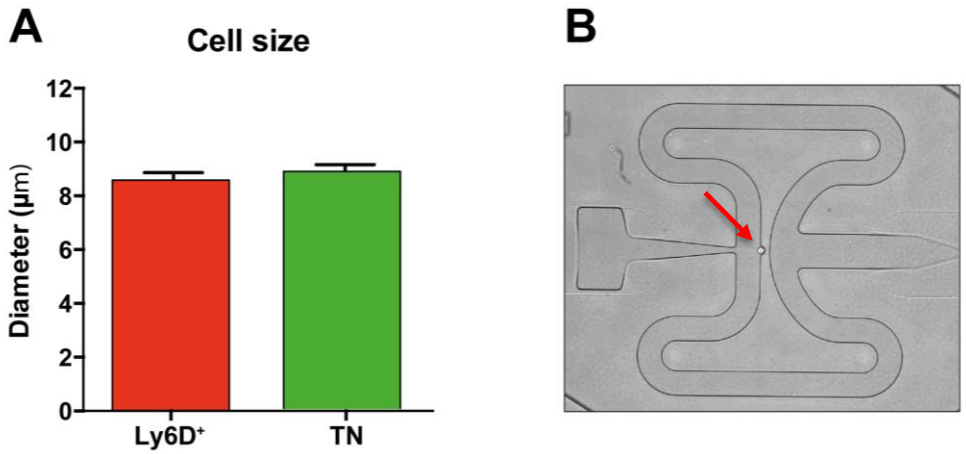
Differential expression analysis was performed using edgeR v3.12.1 [121]. A prior count of 8 was used in order to minimize the large log-fold changes for genes with small number of counts. Genes with a false discovery rate (FDR) <0.05 and abs|$\log_2$(FoldChange)| >1 were considered differentially expressed genes (DEG). For principal component analysis (PCA) average gene expression was centred to zero and only the top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) were used. PCA plots were generated with the ggplot2 v2.1.0 R package. Heatmap with sample pair-wise Pearson's correlation coefficients were also generated with the same subset of genes and visualized with the NMF v0.20.6 R package. Annotated heatmap of gene expression variation of the indicated genes across the samples was also generated with the NMF v0.20.6 R package. Average gene expression was centred to 0 and distances were calculated from Pearson's correlation values of samples using the Euclidean method. MA and Volcano plots were produced using custom R scripts. Gene ontology enrichment analysis was performed with the DAVID v6.7 bioinformatics database, based on Fisher's Exact method [122, 123]. Gene ontology terms of DEG were determined to be significantly enriched when *Pval* <0.05.

5.8 Single-cell RNA sequencing

5.8.1 Capture of single cells

Single cells were captured from *ex-vivo* sorted haematopoietic progenitors on a small-sized (5-10µm cell diameter) C1 Single-Cell Auto Prep IFC for mRNA sequencing (Fluidigm) using the Fluidigm C1 system. Cell diameter of the Ly6D$^+$ and TN cells, imaged on Leica DMI 4000 microscope and measured with ImageJ software, was similar and homogeneous, 8.54 and 8.77µm respectively (**Fig. M3 A**). Therefore, no bias due to cell size or morphology was expected during the capturing. Cells were loaded onto the chip at a concentration of ~300 cells/µl as recommended by the manufacturers and the 96 chambers were inspected by phase-contrast microscopy to determine the number of captured single cells. A total of 3 chips per population were used yielding a good overall capture efficiency, 178 Ly6D$^+$ and 232 TN single

cells (**Fig. M3 B,C**). Subsequently, cells were lysed, the polyA containing mRNA molecules were hybridized to oligo-dT and whole-transcriptome full-length amplified cDNA was prepared by template switching on the C1 fluidigm chip according to the manufacturer's protocol, using the SMARTer Ultra Low RNA kit for the Fluidigm C1 System (Clontech). Quantification of cDNA was done with Quant-iT PicoGreen® dsDNA Assay Kit; TEcan instrument.



C

	TN			Ly6D+		
Total	232			178		
Run	Chip 1	Chip 2	Chip 3	Chip 4	Chip 5	Chip 6
n°	80	77	75	48	69	67
%	83.3	80.2	78.1	50.0	71.9	69.8

Figure M3. Single-cell capturing. (A) Average diameter of Ly6D$^+$ cells (n=24) and TN cells (n=36) determined by phase-contrast microscope and measured with ImageJ. **(B)** Representation of a single-cell captured in one of the 96 chambers of the C1 Fluidigm small chip. Picture taken with phase-contrast microscope in the second chip run of TN cells. **(C)** Capture efficiency. Number (n°) and percentage (%) of single-cells captured per chip (run) or per population (total).

5.8.2 Library preparation

Illumina single-cell libraries were constructed in 96-well plates using the Nextera XT DNA Library Preparation Kit (Illumina) following the protocol supplied by Fluidigm ("Using C1 to Generate Single-Cell cDNA Libraries for mRNA Sequencing"). Briefly, 0.1-0.3ng of harvested cDNA was subjected to tagmentation, a process in which the DNA fragmentation and sequencing adapter ligation occurs in a single step performed by the Nextera XT transposome, followed by purification with AMPure XP beads.

5.8.3 Sequencing

Indexed DNA libraries originated from single cells captured in 3 different chips (288 libraries) were pooled in equal volumes and loaded on one NextSeq 500 High Output flow cell (Illumina). Single-end sequencing was performed on the Illumina NextSeq™ 500 Sequencing System (D-BSSE, Basel) for 76 cycles.

As for the bulk RNA sequencing, reads (76-mers) were received after de-multiplexing in the FastQ format and checked for quality using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, Version 0.11.3). Only the FastQ files corresponding to C1 chambers with a single cell were selected for downstream analysis, thus excluding the ones derived from doublets, debris or empty chambers (previously determined by phase-contrast microscopy). We obtained a total of 360 million reads for the Ly6D$^+$ and 371 million reads for the TN. The average number of reads per cell was 2x10^6 for the Ly6D$^+$ and 1.6x10^6 for the TN (**Fig. M4 A)**.

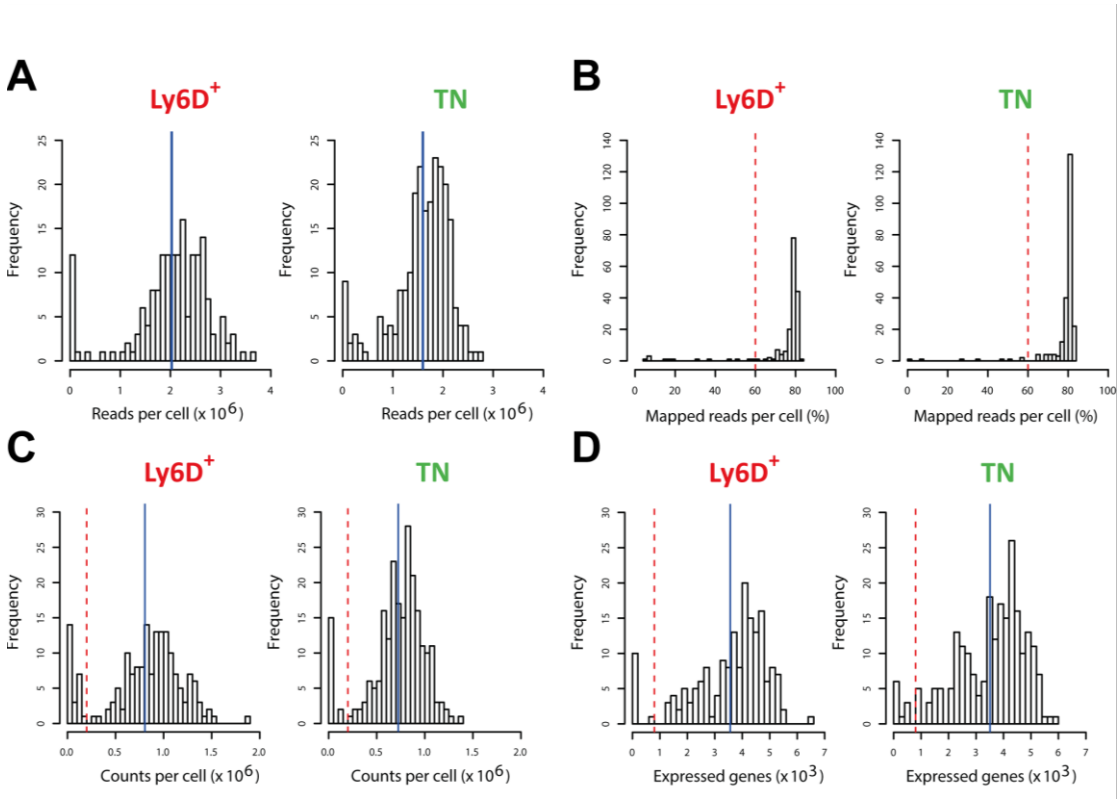


Figure M4. Quantification and quality control of raw sequenced data for Ly6D$^+$ (left panels) and TN (right panels) single cells. Per cell distribution of **(A)** number of sequenced reads, **(B)** percentage of reads mapped to the mouse genome (mm9) out of the total number of reads, **(C)** counts (library size) considering reads mapped to genes (exons only) and, **(D)** total number of genes detected (with at least 1 count). Ly6D$^+$ n=178; TN n=213. Blue line: mean; dotted red line: thresholds applied to the data (any cells not meeting these thresholds failed the quality control and were excluded from the analysis).

5.8.4 Pre-processing of sequencing data

All downstream analysis was performed using the open source R software accessed via RStudio server (R version 3.2.0). Sequencing reads were aligned to the mouse genome assembly, version mm9 (downloaded from UCSC http://genome.ucsc.edu), with SpliceMap [118, 119], included in the R/Bioconductor package QuasR, version 1.10.1. Approximately 80% of total reads were successfully mapped for each sample **(Fig. M4 B).** Subsequently, a count table with gene expression levels was generated using the qCount function from QuasR package and coordinates of RefSeq mRNA genes (http://genome.ucsc.edu, downloaded in December 2013). The expression level was defined as the number of reads that started within any annotated exon of a gene (exon-union model). Total counts per cell were approximately 810,000 for the Ly6D$^+$ and 720,000 for the TN (**Fig. M4 C**), the so-called library size. Genes with no counts across all samples were filtered out from the analysis. At least one read per gene was detected for a total of 14,814 genes across all cells, corresponding to approximately 3,500 expressed genes per cell in both Ly6D$^+$ and TN (**Fig. M4 D**).

During the quality control, cells having less than 60% of mapped reads, less than 200,000 counts, or less than 800 detected genes were filtered out from further analysis (dotted red lines in **Fig. M4**). In total, 365 (152 Ly6D$^+$ and 213 TN) cells out of 410 or 89% passed these criteria.

Raw counts were normalized between cells and genes, expressed as fragments per kilobase of transcript per million mapped reads (FPKM). For visualization purposes, 1 was added to FPKM values and transformed to the log2-scale ($\log_2$FPKM).

5.8.5 Data analysis

If not otherwise specified, the downstream analysis was performed using the 1008 differentially expressed genes (DEG, FDR <0.05 and abs|$\log_2$(FoldChange)| >1) from the bulk RNA-seq experiment when comparing Ly6D$^+$ with TN populations.

Dimensionality reduction was performed with principal component analysis. Average gene expression was centred to zero and PCA plots were generated with the ggplot2 v2.1.0 R package. To visualize the degree of cell-to-cell heterogeneity,

annotated heatmap of sample pair-wise Pearson's correlation coefficients was produced using the NMF v0.20.6 R package. Eight Ly6D$^+$ cells were not considered for subsequent clustering because of their very low transcriptome correlation to any other cell, on average less than 0.3 (**Fig. 11C** left, in results section). Cell clustering was performed using the Partitioning Around Medoids (PAM) method implemented in the cluster v2.0.4 R package [124]. Gene expression was first centered (mean=0) and distances were calculated from cell-to-cell Pearson's correlation values using the Euclidean method. The optimal number of clusters was selected based on silhouette plot, which for Ly6D$^+$ corresponded to K=2 (with average silhouette width of 0.10) and K=3 for the TN (with average silhouette width of 0.13). Cells with negative silhouette width values were excluded while the other 331 cells were assigned to one of the 5 groups. Heatmap with Pearson's correlation coefficients among the clustered groups of cells was generated with the top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) and visualized with the NMF v0.20.6 R package.

Differential gene expression analysis to compare the clustered groups of cells was performed using the 14,528 detected expressed genes across the 331 single cells with edgeR v3.12.1 [121]. A prior count of 0.5 was added to all gene counts in order to minimize the large log-fold changes for genes with small number of counts. Genes with FDR <0.05 and abs|$\log_2$(FoldChange)| >1 were considered as differentially expressed. MA, Volcano, Violin and Scatter plots were produced using custom R scripts.

Gene ontology enrichment analysis was performed with the DAVID v6.7 bioinformatics database, based on Fisher's Exact method [122, 123]. Gene ontology terms of DEGs were determined to be significantly enriched when *Pval* <0.05.

6. Results

6.1 EPLM progenitor population can be divided into at least 4 subpopulations

We have previously characterized an uncommitted and multipotent haematopoietic B220^{+} CD117^{int} CD19^{-} NK1.1^{-} progenitor with combined lymphoid and myeloid potential that we have called "Early Progenitor with Lymphoid and Myeloid potential" (EPLM) [33]. When describing progenitor cells that apparently have multiple lineage potentials, there is a continuing debate concerning whether individual cells within the seemingly phenotypically homogeneous population are truly multipotent or whether the population contains a mixture of cells each with different lineage potentials. With the aim of addressing this question, we decided to examine the expression of additional markers known to be associated with different haematopoietic lineages. These markers were Ly6D, SiglecH and CD11c. Ly6D (Lymphocyte Antigen 6 Complex, Locus D) is a receptor with unknown function used as a specification marker for early B-cell stages. Indeed, a Ly6D^{+} subpopulation of CLP has already been shown to have a B-cell bias [125, 126]. SiglecH is an immunoglobulin-like lectin receptor expressed on mouse plasmacytoid dendritic cells (pDC) that upon binding to sialic acid carbohydrates modulates the secretion of type I interferons and is commonly used as a pDC marker [127, 128]. CD11c (integrin alpha-X) associates with CD18 (integrin beta-2 chain) forming the complex CD11c/CD18 that is the receptor for the complement component iC3b as well as for fibrinogen and is also involved in cell-cell interactions during inflammatory responses. It is important for monocyte adhesion and is used as a murine marker for conventional dendritic cells (cDC) [129].

To test the expression of the selected markers, we combined the reported EPLM gating strategy to define B220^{+} CD117^{int} CD19^{-} NK1.1^{-} EPLM and included Ly6D and SiglecH. This staining combination resulted in the identification of three EPLM fractions (F) both in WT and *Flt3L*tg mice. Thus, F1 was Ly6D^{+} SiglecH^{-}, F2 Ly6D^{+} SiglecH^{+}, and F3 Ly6D^{-} SiglecH^{-} (**Fig. 4A,B** upper panels).

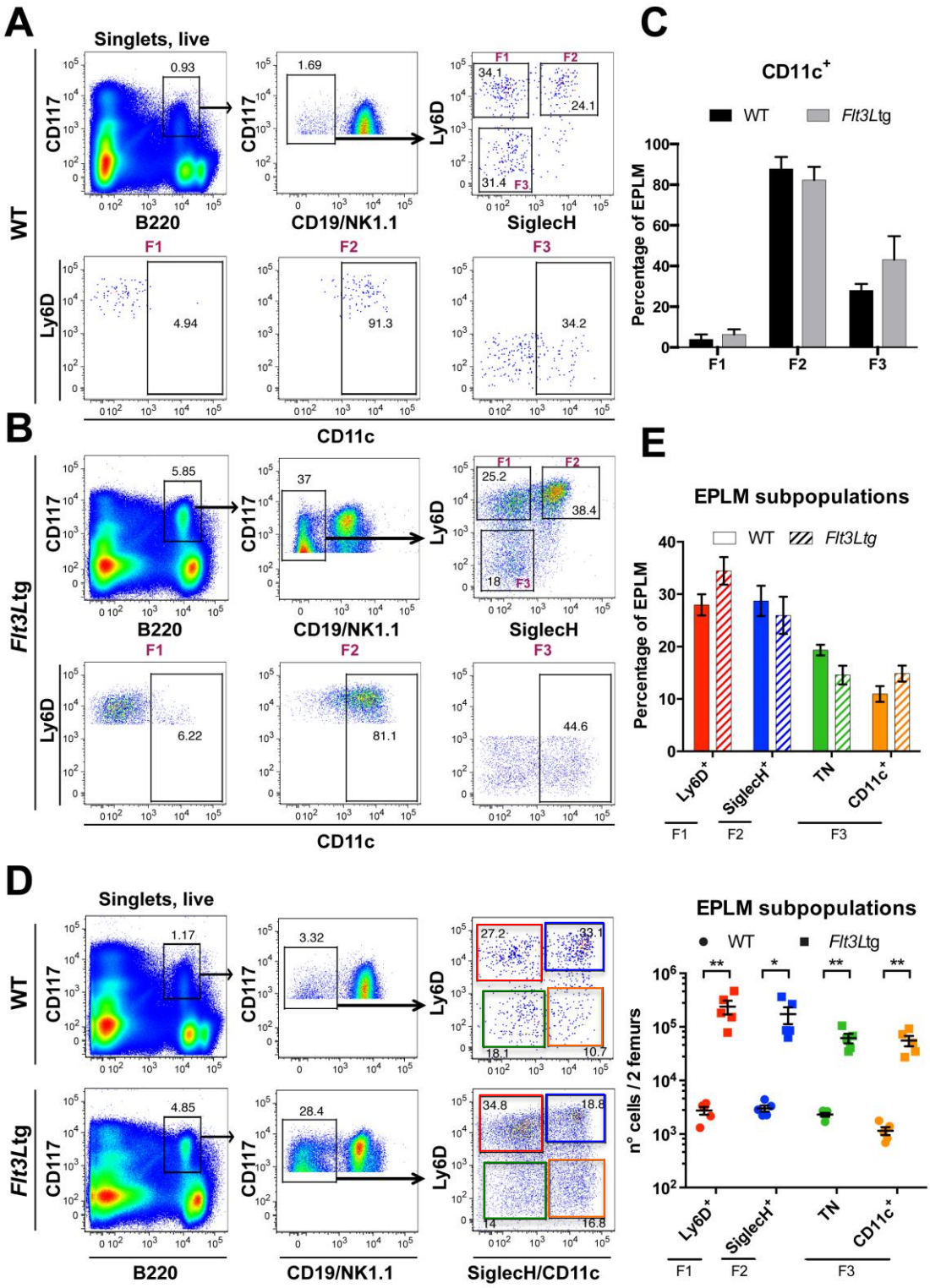


Figure 4. Heterogeneous expression of Ly6D, SiglecH and CD11c on EPLM progenitor population (B220^{+} CD117^{int} CD19^{-} NK1.1^{-}). (A,B) Representative FACS plots of EPLM from the BM (2 femurs) of WT (**A**) and *Flt3L*tg mice (**B**) with the addition of Ly6D and SiglecH identifying three fractions (illustrated as F1, F2 and F3). CD11c expression in each of the fractions is shown in the second row. **(C)** Percentage of CD11c^{+} cells in each EPLM fraction from WT (n=5) and *Flt3L*tg (n=5) mice. Bars show mean±SEM. **(D)** Representative FACS plot showing the gating strategy for the four EPLM subpopulations from WT (upper panel) and *Flt3L*tg (lower panel) mice. **(E)** Summary of EPLM subpopulations as percentages (upper panel) or numbers (lower panel) from WT (n=5) and *Flt3L*tg (n=5) mice. Shown as mean±SEM. Two-tailed unpaired Student *t* tests, *P ≤ 0.05, **P ≤ 0.01.

Subsequently, we explored CD11c expression by these three fractions in both mouse strains. Results indicated that F1 contained only 5.1±0.5% $CD11c^+$ cells whereas F2 were mostly $CD11c^+$ (85.2±2.7%). Interestingly, F3 were heterogeneous for CD11c expression, with about one third (28.1±2.7%) being $CD11c^+$ in WT but considerably more (43.1±9.1%) in *Flt3L*tg mice (**Fig. 4A,B** lower panels and **Fig. 4C**). This result indicates that F3 can be further subdivided into two $CD11c^+$ and $CD11c^-$ fractions resulting in four major EPLM subpopulations. In **Fig. 4D** and in subsequent experiments, we represent the 4 EPLM subpopulations in a simplified manner by staining for Ly6D in one colour and both SiglecH and CD11c using antibodies conjugated with the same fluorochromes in the other. We can thus distinguish the four EPLM subpopulations as: the $Ly6D^+$ $SiglecH^-$ $CD11c^-$ ($Ly6D^+$) single positive subpopulation in red, the $Ly6D^-$ $SiglecH^-$ $CD11c^-$ triple negative (TN) in green, the $Ly6D^+$ $SiglecH^+$ $CD11c^+$ triple positive in blue. Since this is the only subpopulation that contains significant numbers of $SiglecH^+$ cells (**Fig. 4A**,**B** upper right cytogram), they are referred to as $SiglecH^+$. Finally, the $Ly6D^-$ $SiglecH^-$ $CD11c^+$ ($CD11c^+$) subpopulation is shown in orange (**Fig. 4D,E).** In the bone marrow of WT mice, EPLM subpopulations are present in limited numbers ($2.7\pm0.8x10^3$ $Ly6D^+$ cells, $2.3\pm0.7x10^3$ TN cells, $3\pm0.3x10^3$ $SiglecH^+$ cells, and only $1.1\pm0.3x10^3$ $CD11c^+$ cells). In marked contrast, *Flt3L*tg mice contain two orders of magnitude more of each EPLM subpopulation ($2.4\pm0.9x10^5$ $Ly6D^+$, $0.7\pm0.2x10^5$ TN, $1.8\pm1x10^5$ $SiglecH^+$, and only $0.5\pm0.2x10^5$ $CD11c^+$) (**Fig. 4E** lower panel) with minor modifications in their relative frequencies (**Fig. 4E** upper panel). We have previously shown [109] that *Flt3L*tg mice had dramatically increased numbers of EPLM and this result shows that all four subpopulations are affected almost equally.

In conclusion, based on the expression of Ly6D, SiglecH and CD11c, the EPLM progenitor population can be further subdivided into at least four subpopulations whose numbers all increase significantly in *Flt3L*tg mice.

6.2 The heterogeneity of EPLM is reflected in different sets of potentials for each subpopulation

We next wanted to assess if the heterogeneous expression of Ly6D, SiglecH and CD11c cell surface markers by EPLM reflects distinct developmental potentials. Several clonal *in vitro* culture systems exist to quantify the capacity of progenitor cells to differentiate into distinct lineages. These systems use stromal cells and addition of exogenous factors known to support differentiation towards a particular lineage. Stromal cell clones supporting B lymphopoiesis and/or myelopoiesis are well established and extensively used. ST2 stromal cells producing macrophage colony-stimulating factor (M-CSF) support both B lymphopoiesis and/or myelopoiesis although the latter one is favoured [112, 130]. B-cell development from haematopoietic progenitors is much more efficient on OP9 stromal cells [113, 131] because they are derived from the bone marrow of M-CSF deficient *op/op* mice

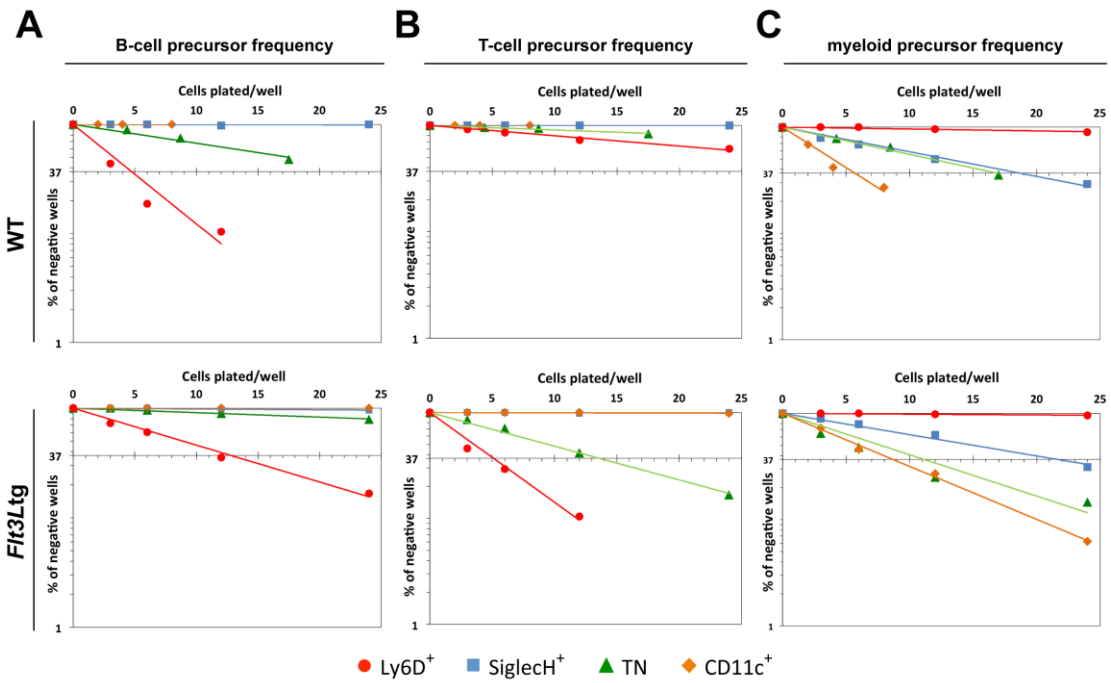


Figure 5. Distinct *in vitro* developmental potentials for the EPLM subpopulations. Limiting dilution analysis of Ly6D$^+$, TN, SiglecH$^+$ and CD11c$^+$ for B-cell **(A)**, T-cell **(B)** or myeloid **(C)** potentials. EPLM subpopulations were sorted from WT (10 pooled mice, upper panels) and *Flt3L*tg (2-pooled mice, lower panels) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7 **(A)**, OP9-DL1 stromal cells in the presence of IL-7 **(B)**, or ST2 stromal cells **(C)**. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope and, when unclear, clone type was confirmed by flow cytometry. TN and CD11c$^+$ subsets from WT were plated at only 3 cell concentrations due to the limited number of cells. The lowest concentration of the two previous subsets was plated in 96 replicates while all the rest in 48. One representative experiment is shown and **Table 2** summarizes the independent experiments performed and indicates the n repetitions for each subpopulation, mice and condition.

[132]. In the absence of M-CSF, macrophage growth is not supported thereby revealing lymphocyte potential [130]. This culture system is normally supplemented with IL-7 and/or other cytokines promoting cell growth such as Flt3L or SCF. Even though B-cell and myeloid propagation was for a long time successfully achieved in monolayer stromal cell cultures, a T-cell system had not been reported. Then, in 2002 Zuniga-Pflücker and colleagues showed for the first time the efficient and long-term commitment and propagation of T cells in a simple stromal cell monolayer system [114, 133]. Use was made of the well-established OP9 stromal cell line, but in addition, since the Delta-Notch system had been identified as instructing T-cell lineage choice [134], OP9 cells were transduced with the Notch-ligand Delta-like 1 (DL1). This generated the OP9-DL1 stromal cell line, which efficiently promoted early stages of T-cell development.

We made use of the above described culture systems and performed a limiting dilution (LD) assay in order to examine the B-cell, T-cell and myeloid precursor frequencies of the four EPLM subpopulations. For that, we sorted them from WT or *Flt3L*tg mice and plated graded number of cells with the appropriate cocktail of stromal cells and cytokines to support B-cell (OP9 + IL7), T-cell (OP9-DL1 + IL7) or myeloid (ST2) differentiation. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope and, when unclear, clone type was confirmed phenotypically by flow cytometry. The result of a representative experiment is shown in **Fig. 5** and a summary of all experiments in **Table 2**. Under B cell conditions, Ly6D$^+$ and TN

		WT			*Flt3L*tg		
		B-cell potential	**T-cell potential**	**Myeloid potential**	**B-cell potential**	**T-cell potential**	**Myeloid potential**
Ly6D$^+$	n	11	6	7	8	4	5
	mean freq	1 in 6.4	1 in 38	< 1 in 99	1 in 11	1 in 4.3	< 1 in 500
SiglecH$^+$	n	4	3	6	3	2	3
	mean freq	< 1 in 500	< 1 in 500	1 in 14	< 1 in 500	< 1 in 500	1 in 30
TN	n	7	3	3	6	3	4
	mean freq	1 in 34	1 in 55	1 in 25	1 in 70	1 in 8.6	1 in 15
CD11c$^+$	n	1	1	1	3	3	3
	mean freq	< 1 in 500	< 1 in 500	1 in 5	< 1 in 500	< 1 in 500	1 in 20

Table 2. Summary table containing mean frequencies of progenitors with B-cell, T-cell and myeloid potential. Number of independent experiments is shown (n) for the different populations.

subpopulations generated colonies, thus revealing B-cell potential. The B-cell precursor frequency was higher in the Ly6D$^+$ cells (about 1 in 5) than in the TN (about 1 in 20) and overall greater for both subpopulations in WT mice compared with their respective counterparts in *Flt3L*tg mice (1 in 12 for Ly6D$^+$ and about 1 in 90 for TN) (**Fig. 5A**). CD11c and SiglecH$^+$ subpopulations did not generate colonies either under B-cell or T-cell conditions (**Fig. 5A,B**). Ly6D$^+$ and TN cells from WT mice generated T cells at low frequencies whereas in *Flt3L*tg mice, frequencies were notably increased (**Fig. 5B**). When EPLM subpopulations were plated on ST2 stromal cells, the Ly6D$^+$ subpopulation did not generate myeloid clones, whereas all other EPLM subpopulations, from either WT or *Flt3L*tg mice, possessed myeloid potentials although at different frequencies (**Fig. 5C**). Therefore, each EPLM subpopulation has different sets of potentials. Whereas Ly6D$^+$ cells are lymphoid restricted, TN show trilineage developmental potential. The SiglecH$^+$ and CD11c$^+$ subpopulations do not show lymphoid potential and as their cell-surface marker profile suggests, they could be (at least a fraction of them) already committed to the pDC and cDC fates respectively. Since in our laboratory we are mostly interested in studying the lymphoid development of EPLM, for the rest of the project we focused on the Ly6D$^+$ and TN subpopulations.

We therefore tested the *in vivo* capacity of Ly6D$^+$ and TN progenitors to reconstitute the B- and T-cell compartments of lymphopenic mice. The two EPLM subpopulations were sorted from WT or *Flt3L*tg mice and equal numbers ($4x10^3$ from WT and $2x10^4$ from *Flt3L*tg) transferred into sub-lethally irradiated *Rag2*-deficient recipient mice [110]. *Rag2*-deficient mice lack mature B and T lymphocytes due to their inability to initiate V(D)J rearrangement. Therefore, the developmental stages observed after antigen receptor rearrangement are generated from donor cells. Flow cytometry of the spleen at 3 weeks following transfer revealed that B-cell compartments were significantly reconstituted in all mice. Both Ly6D$^+$ and TN progenitors either from WT or *Flt3L*tg generated CD19$^+$ IgM$^+$ B cells (**Fig. 6A,C**). Additional analysis of the spleen CD19$^+$ cells revealed the presence of both CD21^{high} CD23$^-$ marginal zone B cells (MZB) and CD21^{int} CD23$^+$ follicular B cells (FB) (**Fig. 6A** lower panels). Therefore, although TN cells present less B-cell precursor frequency *in vitro*, both populations have *in vivo* B-cell developmental potential. Analysis of the

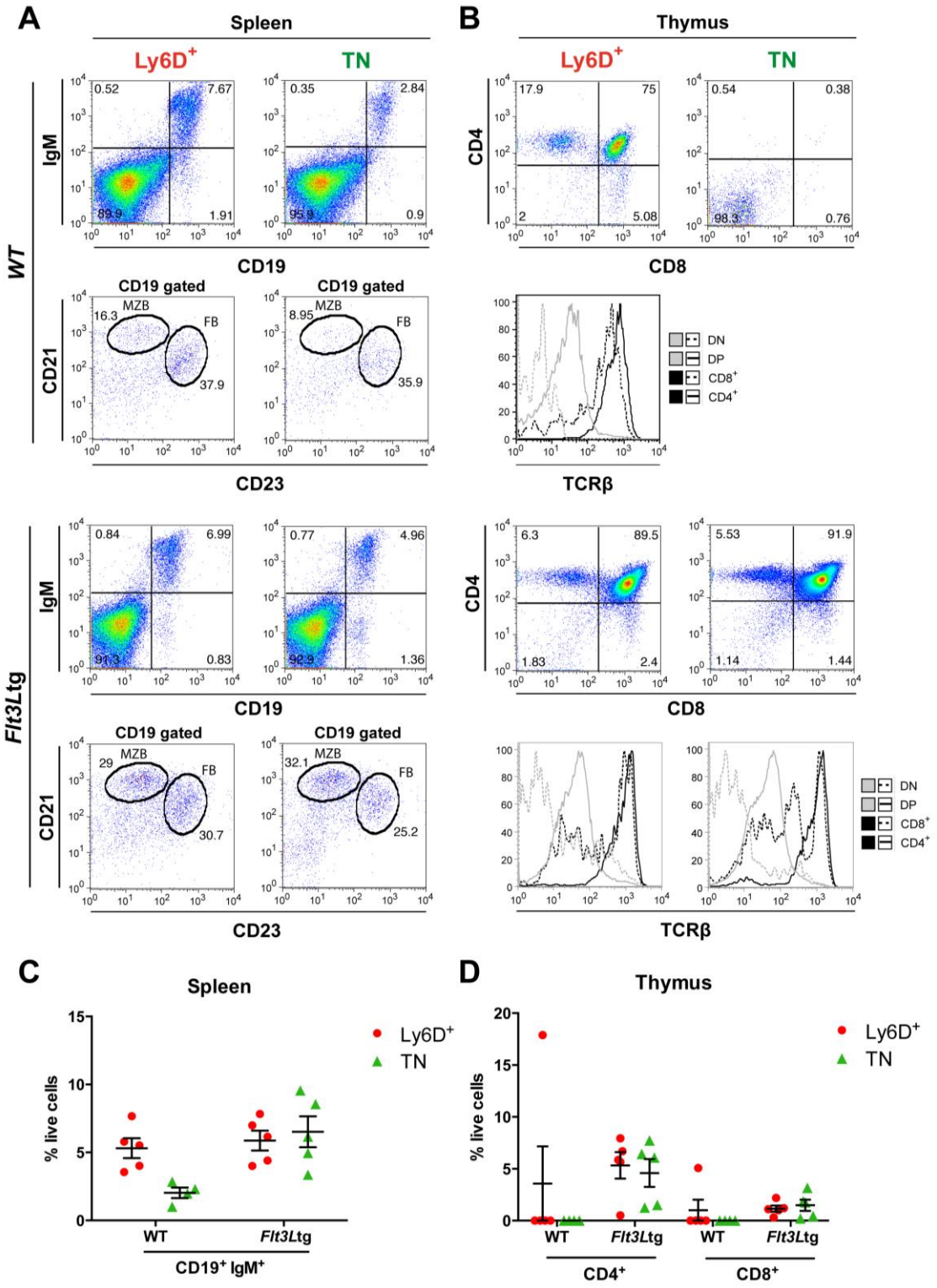


Figure 6. Reconstitution of B- and T-cell compartments in sub-lethally irradiated B6 *Rag2*-deficient mice with Ly6D$^+$ or TN cells from WT or *Flt3L*tg mice. (A) CD19 versus IgM expression on spleen cells 3 weeks after transfer of 4x10^3 WT derived (upper panels) or 2x10^4 *Flt3L*tg derived (lower panels) Ly6D$^+$ and TN subpopulations. Also CD21 and CD23 expression on gated CD19$^+$ spleen cells are shown. **(B)** CD4 and CD8 expression on thymocytes 3 weeks after transfer of 4x10^3 WT derived (upper panels) or 2x10^4 *Flt3L*tg derived (lower panels) Ly6D$^+$ and TN subpopulations. Also TCRβ expression on DN, DP, CD8$^+$ and CD4$^+$ gated thymocytes is shown. **(C,D)** Quantification of CD19$^+$ IgM$^+$ **(C)**, CD4$^+$ and CD8$^+$ **(D)** populations presented as frequency of live cells. Ly6D$^+$ WT (n=5), TN WT (n=4), Ly6D$^+$ *Flt3L*tg (n=5), TN *Flt3L*tg (n=5). Shown as mean±SEM.

thymus showed that, in line with the *in vitro* observations, cells derived from *Flt3L*tg were much more efficient at reconstituting the thymus than those from WT mice (**Fig. 6B,D**). Thus, WT derived progenitors could only reconstitute the thymus of one out of nine mice whereas progenitors from *Flt3L*tg mice reconstituted seven out of ten animals (**Fig. 6D**). Comparing the two EPLM subpopulations, in WT mice only the Ly6D$^+$ subpopulation had any T cell reconstitution potential (1/5 vs 0/5) whereas in *Flt3L*tg mice, both subpopulations had quite robust reconstitution potential (4/5 vs 3/5) (**Fig. 6D**). In the T-cell reconstituted thymus, we observed normal levels of CD4$^+$ CD8$^+$ double positive (DP), and CD4$^+$ or CD8$^+$ single positive T cells. Simultaneous staining for TCRβ indicated that, like in the normal thymus, TCRβ expression on DP was lower compared to that on CD4$^+$ cells, and that the CD8$^+$ population contained a mixture of TCRβ$^+$ and TCRβ^{low} "immature" CD8$^+$ cells (**Fig. 6B**). Moreover, CD4$^-$ CD8$^-$ double negative (DN) thymocytes were largely TCRβ$^-$ and only a small fraction was TCRβ$^+$. Interestingly, although Ly6D$^+$ *Flt3L*tg cells showed higher *in vitro* T cell potential (**Fig. 6A,B** lower panels), they reconstituted better the B-cell than the T-cell compartment.

6.3 Ly6D$^+$ EPLM has a lymphoid whereas TN a myeloid genetic signature

In order to characterize EPLM subpopulations at the molecular level, we performed gene expression profiling (RNA-seq) of bulk populations of Ly6D$^+$ and TN EPLM subpopulations as well as CD117$^+$ CD19$^+$ Pro-B cells from *Flt3L*tg mice. We included the latter population as an already B-cell lineage committed bone marrow population and thus downstream of EPLMs. Ly6D$^+$ and TN were sorted as indicated in **Fig. 4D**. All populations were sorted four separate times from the pooled bone marrow of two 6-8 week-old male *Flt3L*tg mice, in which EPLM numbers are significantly increased. All samples were processed as explained in Materials and Methods and all passed the quality control (**Fig. M1 and Fig. M2**). Moreover, all biological replicates presented high transcriptome correlation (> r=0.990) (**Fig. 7**). Therefore, we proceeded to apply Principal Component Analysis (PCA) to the data

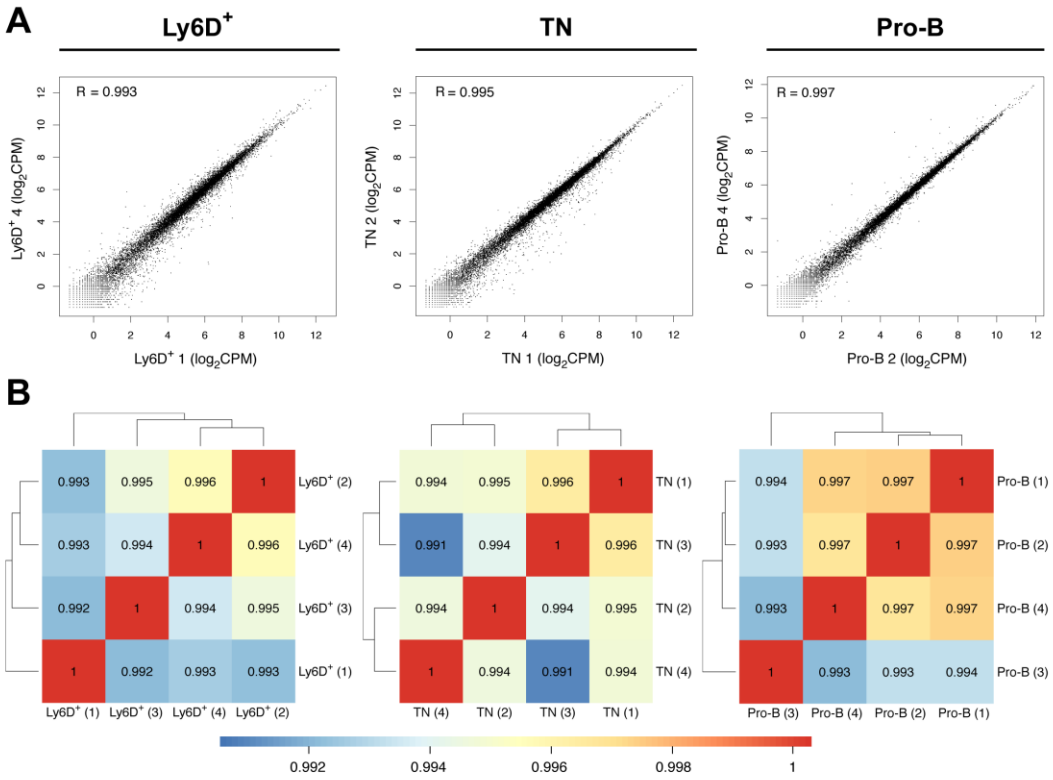


Figure 7. Quality of the Ly6D$^+$, TN and Pro-B replicates. (A) Scatterplots with individual expression profiles of two representative biological replicates for each population**. (B)** Heatmaps with pair-wise Pearson's transcriptome correlation per population.

and calculated sample pair-wise Pearson's correlations. The highest variation among samples (71.13% reflected in PC1 axis) accounted for their developmental stage; with the uncommitted EPLM subpopulation on the left and the committed Pro-B cells on the right (**Fig. 8A**). Pair-wise transcriptome correlation revealed that, in line with PCA, Ly6D$^+$ and TN were the two populations with the highest transcriptome association (r=0.973) whereas from the two EPLM subpopulations, the Ly6D$^+$ subset was the closest to Pro-B cells (r=0.918 Ly6D$^+$/Pro-B versus r=0.886 TN/Pro-B) and therefore to the B-cell lineage (**Fig. 8B**).

We further explored transcriptome differences by performing differential expression analysis and this is summarised in the table in **Fig. 8C** showing the total number of Differentially Expressed Genes (DEG) for each pair-wise comparison as well as the number and the fraction corresponding to up-regulated and down-regulated genes. To be considered as DEG, the gene expression had to be at least two times higher or lower in one population compared with the other and this change in expression had to be significant (corrected p-value <0.05). A considerable

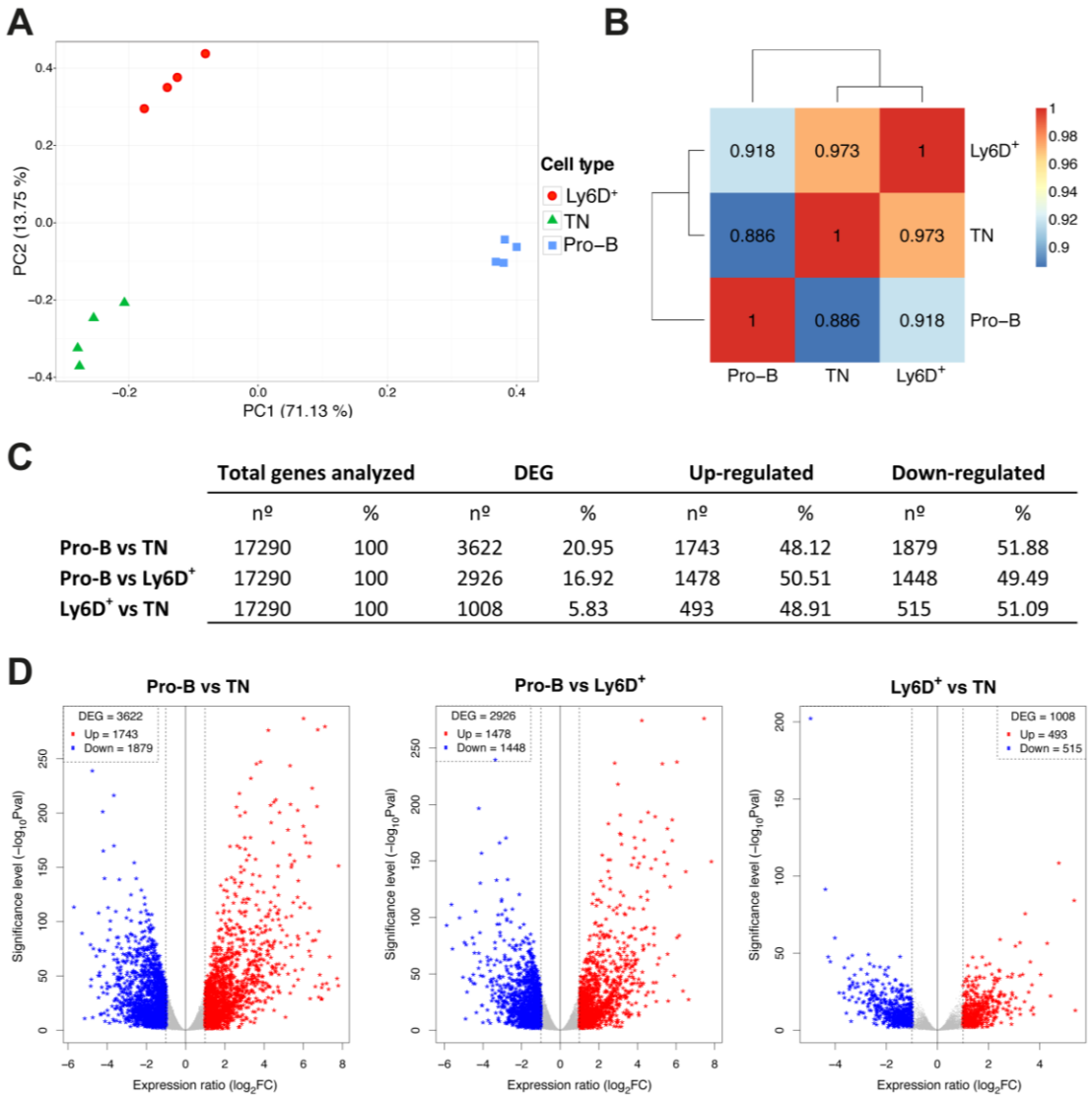


Figure 8. Of the two EPLM subpopulations, Ly6D$^+$ cells are closer to the Pro-B cells. (A) Principal component analysis and **(B)** heatmap with pair-wise Pearson's transcriptome correlation of Ly6D$^+$, TN and Pro-B replicates and averaged populations respectively**.** The top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) were used. **(C)** Summary table of differential expression analysis containing the differentially expressed, up-regulated and down-regulated genes in number and percentage of each pair-wise transcriptome comparison. Provided excel file 1 **(1_DEGlists_bulkRNAseq)** contains the complete lists of DEG. **(D)** Volcano plot (plotted significance against expression ratio) of each pair-wise transcriptome comparison. Each dot/star represents a gene. Grey dots: not DEGs; red star: up-regulated genes; blue stars: down-regulated genes.

fraction (21% and 17%) of genes was differentially expressed when comparing Pro-B with either Ly6D$^+$ or TN subpopulations respectively whereas only about 6% (1008 genes) presented a significant change in expression between the two EPLM subpopulations (**Fig. 8C**), thus suggesting again that Ly6D$^+$ and TN cells are more related to each other than to Pro-B cells. In addition, the graphical representation of the differential expression analysis (volcano plots, **Fig. 8D**) also revealed that the fold changes and significances of DEGs in the Ly6D$^+$vsTN was less (**Fig. 8D** right panel)

than when they were individually compared with Pro-B (**Fig. 8D** left and middle panels). Interestingly, there was a similar fraction of up-regulated and down-regulated genes in each comparison, indicating no predominant activation or repression of genetic programmes from one haematopoietic stage to the other (**Fig. 8C**).

Up-regulated genes

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0046649~lymphocyte activation	19	3.81E-07	4.35
GOTERM_BP_FAT	GO:0042113~B cell activation	10	6.83E-05	5.60
GOTERM_BP_FAT	GO:0042100~B cell proliferation	5	1.14E-04	18.20
GOTERM_BP_FAT	GO:0050851~antigen receptor-mediated signaling pathway	7	1.47E-04	8.50
GOTERM_BP_FAT	GO:0030098~lymphocyte differentiation	11	2.77E-04	4.22
GOTERM_BP_FAT	GO:0033151~V(D)J recombination	4	8.92E-04	19.42
GOTERM_BP_FAT	GO:0042110~T cell activation	10	0.0013	3.77
GOTERM_BP_FAT	GO:0030217~T cell differentiation	8	0.0017	4.60
GOTERM_BP_FAT	GO:0046651~lymphocyte proliferation	6	0.00255	6.24
GOTERM_BP_FAT	GO:0050853~B cell receptor signaling pathway	4	0.0036	12.48
GOTERM_BP_FAT	GO:0016444~somatic cell DNA recombination	4	0.0131	7.94
GOTERM_BP_FAT	GO:0045165~cell fate commitment	9	0.0195	2.67
GOTERM_BP_FAT	GO:0030183~B cell differentiation	5	0.0204	4.75
GOTERM_BP_FAT	GO:0002377~immunoglobulin production	4	0.0303	5.83

Down-regulated genes

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0009611~response to wounding	41	1.91E-14	4.22
GOTERM_BP_FAT	GO:0006954~inflammatory response	30	5.52E-12	4.77
GOTERM_BP_FAT	GO:0006909~phagocytosis	12	7.74E-08	8.76
GOTERM_BP_FAT	GO:0002274~myeloid leukocyte activation	9	4.14E-06	9.19
GOTERM_BP_FAT	GO:0032680~regulation of TNF production	8	6.81E-06	10.59
GOTERM_BP_FAT	GO:0006897~endocytosis	18	2.06E-05	3.42
GOTERM_BP_FAT	GO:0045087~innate immune response	13	4.25E-05	4.34
GOTERM_BP_FAT	GO:0016064~immunoglobulin mediated immune response	9	3.45E-04	5.11
GOTERM_BP_FAT	GO:0042742~defense response to bacterium	11	8.64E-04	3.64
GOTERM_BP_FAT	GO:0030099~myeloid cell differentiation	10	0.0014	3.72
GOTERM_BP_FAT	GO:0006957~complement activation, alternative pathway	4	0.0022	14.30
GOTERM_BP_FAT	GO:0019882~antigen processing and presentation	9	0.0029	3.70
GOTERM_BP_FAT	GO:0030593~neutrophil chemotaxis	4	0.0109	8.41
GOTERM_BP_FAT	GO:0042116~macrophage activation	3	0.0301	10.73

Table 3. Selection of enriched Biological Processes (eBP) in up-regulated and down-regulated genes of the Ly6D$^+$ versus TN comparison. Complete list in excel file 2 (**2_eBP_Ly6DvsTN**).

We next investigated in detail the genetic differences between the apparently transcriptomically related Ly6D$^+$ and TN EPLM subpopulations. For this, we studied the nature of the DEG (Ly6D$^+$ vs TN) by gene ontology (GO) enrichment analysis. This analysis revealed that the 500 up-regulated genes in Ly6D$^+$ cells were enriched for lymphoid biological processes such as activation, proliferation and differentiation of B and T cells, VDJ recombination and immunoglobulin production… (**Table 3**, up). Moreover, the genes that accounted for B-cell related biological

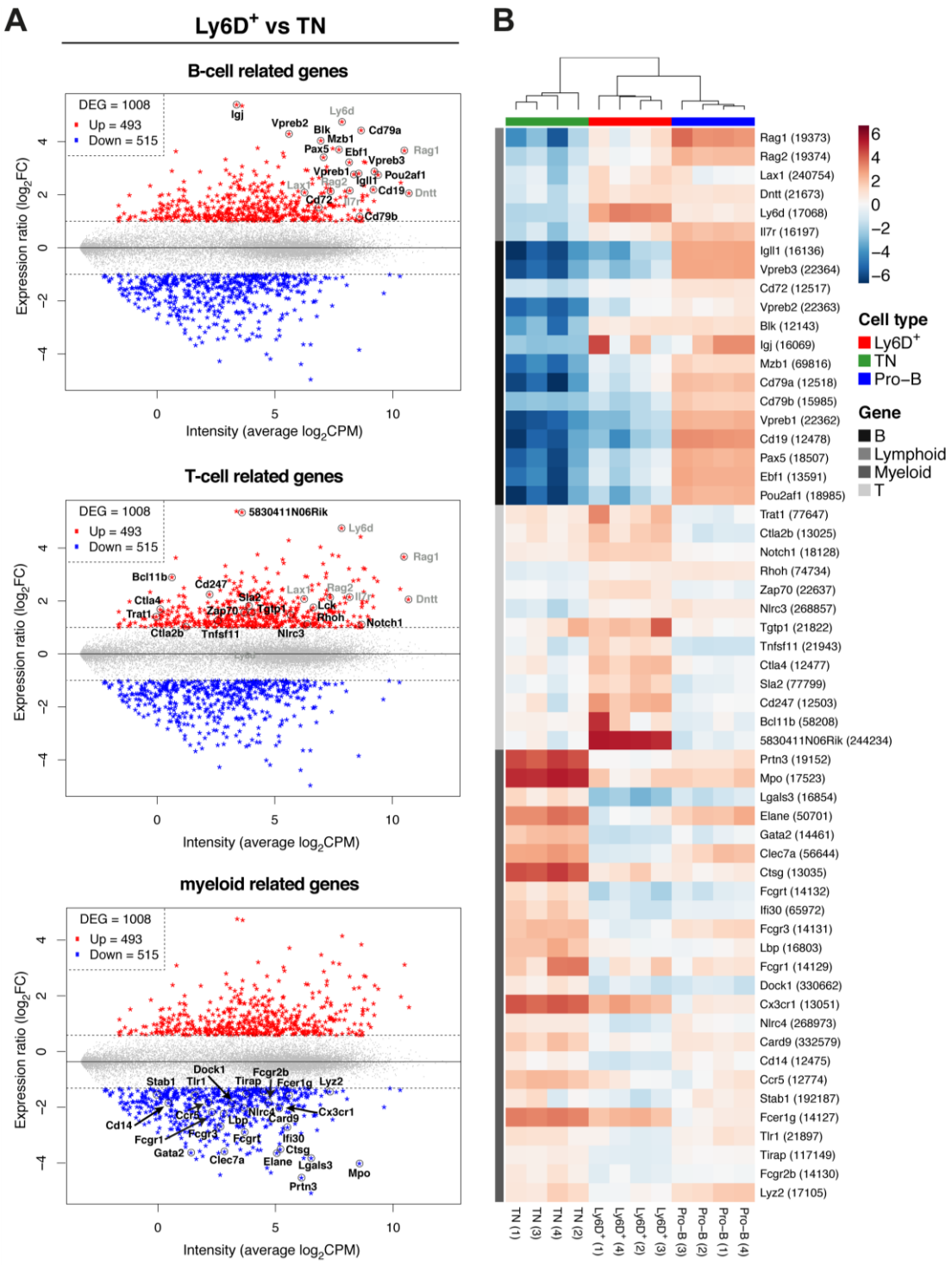


Figure 9. Ly6D$^+$ and TN EPLM subpopulations present distinct genetic signatures. (A) MA plot (plotted expression ratio against average expression intensity) of Ly6D$^+$ vs TN transcriptome comparison. B-cell (upper panel), T-cell (middle panel) and myeloid (lower panel) related genes are indicated. Grey labels correspond to characteristic genes of both B and T cells. **(B)** Heatmap of expression of lineage-specific genes indicated in **A.** The colour gradient illustrates the centred gene expression level (log_2CPM, mean=0 across all samples and per gene)**.**

processes were highly expressed and most of them within the top 50 of up-regulated genes (**Fig. 9A**, upper panel). Among these, were genes encoding the B-cell related transcription factors *Pax5*, *Ebf1 and Pou2af1* (*Obf1*), the recombinase machinery

Rag1 and *Rag2,* the surrogate light chains *VpreB1*, *VpreB2*, *VpreB3* and *Igll1* (lambda 5) of the pre B-cell receptor (pre-BCR), the signalling immunoglobulin α (*Cd79a*) and β (*Cd79b*) chains of the pre-BCR complex, the non-receptor tyrosine kinase *Blk* involved in BCR signaling, the receptor for interleukin-7 *Il7r* and other lymphoid related genes (*Dntt*, *Lax* and, as expected, *Ly6d* itself). Interestingly, although $Ly6D^+$ cells were sorted as $CD19^-$ cells, mRNA expression of the B cell co-receptor *CD19* was already detected. Taken together, these results suggest that qualitatively, $Ly6D^+$ cells express a B-cell genetic signature characteristic of $CD19^+$ Pro-B cell stage. However, quantitatively, the overall expression of these genes is markedly lower than Pro-B cells (**Fig. 9B**). In contrast to B-cell related genes, the genes accounting for T-cell related biological processes in the 500 up-regulated genes in $Ly6D^+$ cells presented lower expression ratios and variable expression intensities overall (**Fig. 9A**, middle panel). Among these genes were the T-cell transcription factor *Bcl11b*, the *Notch1* receptor, a master regulator of T-cell commitment whose signaling represses the expression of genes related with other lineages, the signalling CD3 zeta chain (*Cd247)* of the T-cell receptor (TCR) complex, genes involved in pre-TCR signalling such as *Lck* (non-receptor tyrosine kinase) *Rhoh* (related GTP-binding protein) *Zap70* (tyrosine kinase) and *Sla2* (Src-like-adapter protein), *Trat1* (an adaptor protein that stabilizes the TCR/CD3 complex at the surface of T-cells), the *tnfsf11* cytokine and *Nlrc3* (positive regulators of T cell activation), and the inhibitory T-cell related receptors *Ctla4* and *Ctla2b.* This T-cell genetic signature is exclusive to the $Ly6D^+$ subpopulation (**Fig. 9B**) and is the "feature" that separates them from both the Pro-B and TN cells along the PC2 of the principal component analysis (**Fig. 8A)**.

Analysis of the 500 down-regulated genes revealed that they were largely related with myeloid and innate biological processes such as inflammation, phagocytosis, responses to bacteria, yeast and fungi, and macrophage activation (**Table 3**, down). Some key genes accounting for these processes were *Mpo*, *Elane* and *Ctsg* (enzymes with microbicidal activity), *Prtn3* (a serine protease that degrades elastin, fibronectin, laminin, vitronectin, and collagen), the phagocytic Fc receptors *Fcgr2b*, *Fcgrt*, *Fcer1g* and *Fcgr1,* the pathogen recognition receptor *Tlr1* (Toll-like receptor 1)*, Gata2* (transcriptional regulator of phagocytosis), *Clec7a* (involved in TLR2-mediated inflammatory responses), the polysaccharide binding protein *Lbp* and

the chemokine receptor *Cx3cr1* involved in myeloid leukocyte activation (**Fig. 9A**, lower panel, and **B**). Therefore, TN cells would appear to present a myeloid genetic signature.

From this transcriptomic analysis, we conclude that i) EPLM subpopulations are distinct from one another and are both distinct from Pro-B cells ii) of the two EPLM subsets, Ly6D$^+$ cells are closer to Pro-B cells iii) whereas the Ly6D$^+$ subset has a largely lymphoid genetic signature, that of the TN subset is more myeloid.

6.4 EPLM subpopulations are developmentally related

We have seen that Ly6D$^+$ and TN cells have distinct genetic signatures and different *in vitro* developmental potentials. However, they are fractions of the original EPLM population. The question arose whether as a population EPLM are composed of developmentally unrelated fractions or whether there is a precursor-product relationship between them. To address this question, we sorted the two EPLM subpopulations from *Flt3L*tg mice and cultured them in the presence of Flt3L and IL-7. An initial number of 5x10^4 cells were plated and from day 2 to day 6 we monitored the emergence of Ly6D$^+$ and CD19$^+$ cells by flow cytometry. We observed that, already after 48 hours, almost half of the cells sorted as Ly6D negative (TN), up-regulated Ly6D (**Fig. 10A** lower left panel). This indicates that the TN EPLM

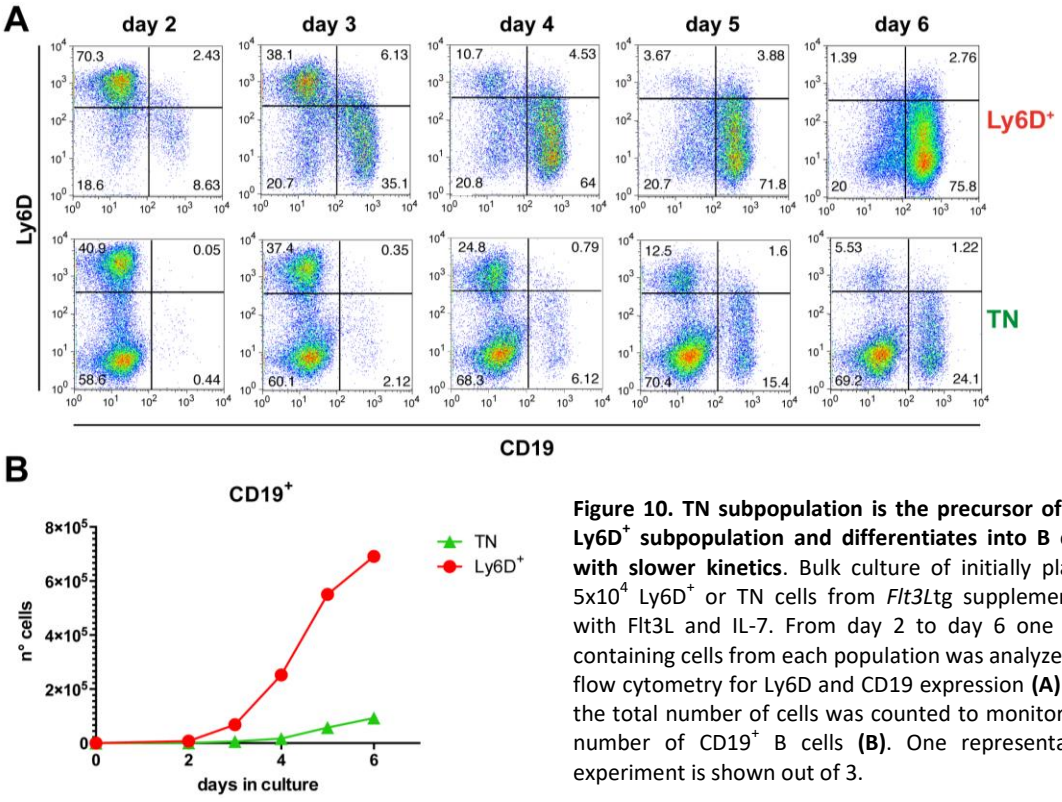


Figure 10. TN subpopulation is the precursor of the Ly6D$^+$ subpopulation and differentiates into B cells with slower kinetics. Bulk culture of initially plated 5x10^4 Ly6D$^+$ or TN cells from *Flt3L*tg supplemented with Flt3L and IL-7. From day 2 to day 6 one well containing cells from each population was analyzed by flow cytometry for Ly6D and CD19 expression **(A)** and the total number of cells was counted to monitor the number of CD19$^+$ B cells **(B)**. One representative experiment is shown out of 3.

subpopulation can give rise to the Ly6D$^+$ subset. B cells expressing CD19 were detected around day 4 reaching 24% of cells at day 6, whereas by this time, Ly6D expression had decreased (**Fig. 10A** lower panels). In contrast, Ly6D$^+$ cells had differentiated into B cells already by day 2 and 75% were CD19$^+$ at day 6 (**Fig. 10A** upper panels). Moreover, the total number of B cells generated from Ly6D$^+$ was dramatically higher than those generated from TN cells (**Fig. 10B**). These results indicated that, as observed in the limiting dilution assays, Ly6D$^+$ cells have higher B-cell precursor frequency and differentiate into B cells with faster kinetics compared with TN. The slower kinetics observed in the TN cells indicate that the TN are the direct precursors of the Ly6D$^+$ and, therefore, they might first move to the Ly6D$^+$ stage before differentiating into B cells.

6.5 Single-cell RNA sequencing reveals distinct degree of molecular heterogeneity of Ly6D$^+$ and TN EPLM subpopulations

As a population, the TN subset of EPLM would appear to have multilineage developmental potential **(Fig. 5)** prompting the question whether they are composed of a mixture of cells. If this were the case, which fraction of TN EPLM constitutes the precursor of the Ly6D$^+$ fraction? Therefore, we further explored the heterogeneity of EPLM subpopulations by performing single-cell RNA sequencing. Ly6D$^+$ (B220$^+$ CD117^{int} CD19$^-$ NK1.1$^-$ Ly6D$^+$ SiglecH$^-$ CD11c$^-$) and TN (B220$^+$ CD117^{int} CD19$^-$ NK1.1$^-$ Ly6D$^-$ SiglecH$^-$ CD11c$^-$) were sorted from the same *Flt3L*tg mice used for the bulk RNA sequencing and utilized for the capture of single cells with the C1 Fluidigm system. A total of 178 Ly6D$^+$ and 232 TN single cells were captured (**Fig. M3 C**) and their subsequent library preparation and sequencing were performed as explained in Materials and Methods. During the quality control of the sequenced data only cells with more than 60% of mapped reads, at least 2x10^5 counts and more than 800 detected genes were selected for further analysis, resulting in 152 Ly6D$^+$ and 213 TN single cells (**Fig. M4**). In Principal Component Analysis cells did not cluster according to the chip they were captured, thus revealing that there was no batch effect (**Fig. 11A).** Instead, in the same PCA with the cells coloured according

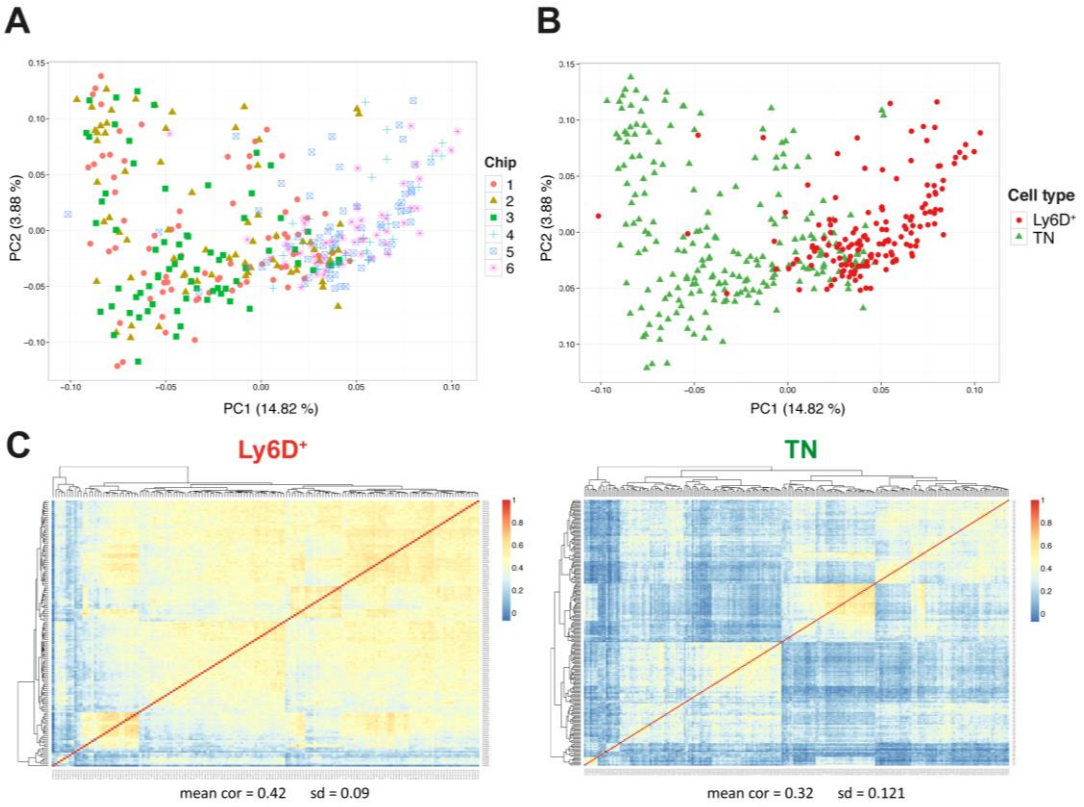


Figure 11. Transcriptomic heterogeneity of TN and Ly6D$^+$ EPLM subpopulations by single-cell RNA sequencing. (A,B) Principal component analysis of 152 Ly6D$^+$ and 213 TN single cells using as gene set the 1008 differentially expressed genes from the bulk RNA-seq experiment when comparing Ly6D$^+$ with TN populations. Cells are coloured according to the chip they were captured **(A)** or the cell type **(B)**. Dynamic 3D PCA plots are provided in the attached documents: **3D_PCAplot_1** and **3D_PCAplot_2**. **(C)** Heatmap with cell-to-cell Pearson's transcriptome correlation of Ly6D$^+$ (left) and TN (right) cells calculated with the same gene set as for the PCA. Mean correlation value (mean cor) and standard deviation (sd) is shown.

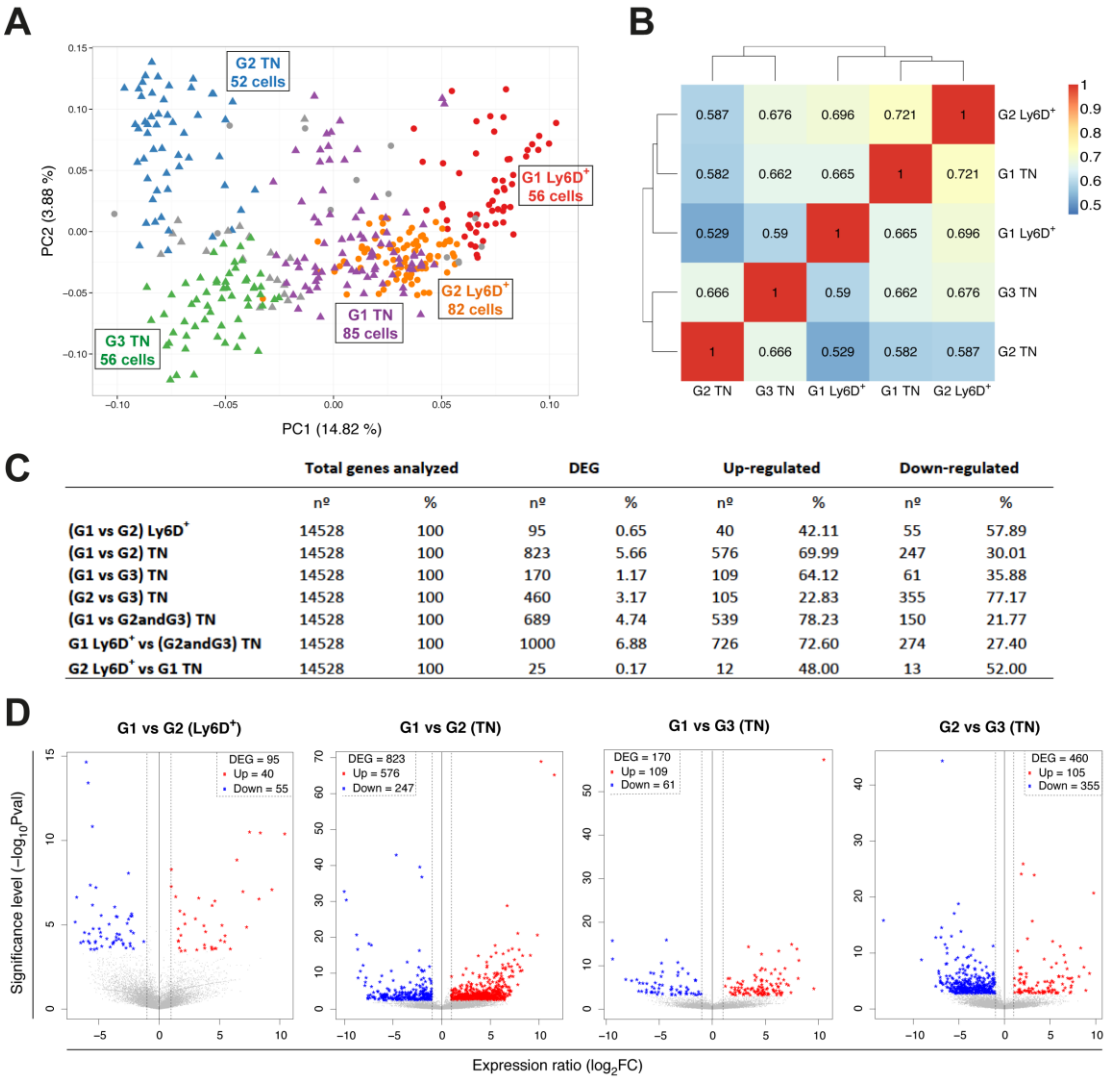
to their cell type, we found that the first component (PC1 axis) partially segregated the TN (on the left) from the Ly6D$^+$ (on the right) with some cells overlapping in the middle (**Fig. 11B** and attached **3D_PCAplot_2**), suggesting that the two EPLM subpopulations, or at least a fraction, are transcriptionally related. PC2 (**Fig. 11B**) and PC3 (attached **3D_PCAplot_2**) showed that Ly6D$^+$ cells are distributed along a single pathway that is relatively homogeneous whereas the TN cells are more differently distributed and exhibit a branching pattern, thereby indicating that they might be more heterogeneous. In order to better address this, we quantified the degree of cell-to-cell heterogeneity by calculating the cell pair-wise Pearson's correlation coefficients and generating a correlation heatmap for each cell type (**Fig. 11C**). Ly6D$^+$ single cells showed an overall higher transcriptome correlation (predominant yellow/orange colour and a mean correlation value of 0.42, **Fig. 11C** left) compared with the TN single cells (predominant blue colour and a mean correlation value of

0.32, **Fig. 11C** right). Moreover, Ly6D$^+$ cells showed a seemingly homogeneous correlation whereas the TN cells presented clusters of cells transcriptomically related to each other (yellow) but very different to the rest (blue). These results indicate that the TN EPLM subpopulation has in turn a heterogeneous transcriptome and might be composed by a mixture of cells.

6.6 Identification of two Ly6D$^+$ and three TN subgroups with distinct genetic signatures

The next question after identifying a heterogeneous subset was to unravel how many subgroups did it contain. Therefore, we performed cell clustering using the Partioning Around Medoids (PAM) method as explained in Materials and Methods. The analysis revealed two robust groups (G1 and G2) of Ly6D$^+$ and three (G1, G2 and G3) of TN cells as illustrated in the PCA plot coloured according to the subgroups (**Fig. 12A** and attached **3D_PCAplot_3**). Thus, the Ly6D$^+$ population is further subdivided into G1 Ly6D$^+$, composed of 56 cells (red) and G2 Ly6D$^+$, composed of 82 cells (orange), whereas the TN population is subdivided into G1 TN with 85 cells (purple), G2 TN composed of 52 cells (blue) and G3 TN with 56 cells (green). In order to explore the degree of similarity or divergence among the clustered groups of cells we calculated their pair-wise transcriptome correlation, which revealed that the two Ly6D$^+$ subgroups had higher transcriptome association (R=0.696) than those observed between any of the TN subgroups (R=0.582 G1/G2, R=0.662 G1/G3, R=0.666 G2/G3) (**Fig. 12B**). Interestingly, the two groups of cells with the highest transcriptome correlation are part of different EPLM subpopulations, namely the G2 Ly6D$^+$ and G1 TN with R=0.721 **(Fig. 12B)**, thereby suggesting that the G1 TN (purple cells in **Fig. 12A**) group could be the fraction of the TN population that, as we have observed in culture, is the precursor of the Ly6D$^+$ cells (circles in **Fig. 12A**). Finally, the subgroup having the most distinct transcriptome profile is the G2 TN (**Fig. 12B** left column, and blue cells in **Fig.12A**).

We further studied genetic differences by applying differential expression analysis to the clustered groups of cells. The table in **Fig. 12C** shows all the



C

	Total genes analyzed		DEG		Up-regulated		Down-regulated	
	nº	%	nº	%	nº	%	nº	%
(G1 vs G2) Ly6D$^+$	14528	100	95	0.65	40	42.11	55	57.89
(G1 vs G2) TN	14528	100	823	5.66	576	69.99	247	30.01
(G1 vs G3) TN	14528	100	170	1.17	109	64.12	61	35.88
(G2 vs G3) TN	14528	100	460	3.17	105	22.83	355	77.17
(G1 vs G2andG3) TN	14528	100	689	4.74	539	78.23	150	21.77
G1 Ly6D$^+$ vs (G2andG3) TN	14528	100	1000	6.88	726	72.60	274	27.40
G2 Ly6D$^+$ vs G1 TN	14528	100	25	0.17	12	48.00	13	52.00

Figure 12. Cell clustering identifies three TN and two Ly6D$^+$ distinct subgroups. (A) PCA with shape according to cell type (circles Ly6D$^+$, triangles TN) and colour according to the subgroups revealed by PAM clustering method (materials and methods). Text boxes indicate the group name and the number of cells. Grey: excluded cells due to low correlation or negative width silhouette values as explained in materials and methods. **3D_PCAplot_3** attached**. (B)** Heatmap with pair-wise Pearson's transcriptome correlation of Ly6D$^+$ and TN subgroups. Average expression across all detected genes was calculated for each of the five cell clusters, and the top 50% of genes with highest variance across analysed dataset (calculated as inter-quartile range) were used. **(C)** Summary table of differential expression analysis containing the differentially expressed, up-regulated and down-regulated genes in number and percentage of the indicated transcriptome comparisons. Provided excel file **(3_DEGlists_scRNAseq)** contains the complete lists of DEG. **(D)** Volcano plot (plotted significance against expression ratio) of the indicated pair-wise transcriptome comparisons. Each dot/star represents a gene. Grey dots: not DEGs; red star: up-regulated genes; blue stars: down-regulated genes.

comparisons analysed reporting the number and percentage of DEG (FDR <0.05 and abs|$\log_2$(FoldChange)| >1), up-regulated and down-regulated genes. The complete table containing the DEG for each comparison is provided as a supplementary file (**3_DEGlists_scRNAseq**). Less than 1% of the genes were differentially expressed when comparing the two Ly6D$^+$ subgroups (**Fig. 12C** first row) and with overall low significance level (**Fig. 12D**, left). In contrast, comparisons among the TN subgroups

yielded more differentially expressed genes (1-6%, **Fig. 12C** second to fourth rows), and with overall higher significance levels and fold ratios (**Fig. 12D**). Of note, when comparing the two subgroups with the highest transcriptome correlation, namely G2 Ly6D$^+$ and G1 TN, only 25 genes were differentially expressed with *Ly6d*, as expected, on the top (complete DEG lists in **3_DEGlists_scRNAseq** excel file), therefore confirming that these two subgroups of cells sorted as two phenotypically distinct EPLM subpopulations (Ly6D$^+$ and Ly6D$^-$) are related.

We next performed a detailed screening based on the DEG lists in order to unravel the gene expression patterns of each subgroup. **Fig. 13 (A-C)** shows a collection of genes exclusively or more highly expressed in one of the clustered groups of cells. The G1 Ly6D$^+$ red subgroup up-regulates genes (**Fig. 13A**) related with B-cell biological processes (**Table 4A**). Although we have previously shown that the entire Ly6D$^+$ population is lymphoid restricted (**Fig. 5**) and has a strong B-cell genetic signature (**Fig. 9A** upper panel), single-cell transcriptomic analysis reveals that this signature is mostly contained within the red subgroup, as is exemplified by *CD19* expression **(Fig. 13A)**, a hallmark of B-lymphopoiesis. The G2 TN blue cluster of cells up-regulates genes characteristic of the cDC lineage such as the genes encoding the α and β chains of the major histocompatibility complex (MHC) class II (*H2-Aa*, *H2-Ab1*, *H2-Eb1*), the MHC class II-associated invariant chain (Ii, *Cd74*), the transcriptional co-activator of MHC class II genes (*Ciita*), the E3 ligase (*March1*) that ubiquitinates the cytoplasmic tail of MHC class II regulating their steady-state level of expression, the transcriptional regulators *Id2* and *Batf3* (expressed by cDC and necessary for the development of CD8α cDC) and the integrin *Itgb7* related with leukocyte migration to mucosal tissues (**Fig. 13B**). Apart from antigen processing and presentation biological processes, the genes up-regulated in the G2 TN cells are also involved in actin cytoskeleton organization, leukocyte adhesion, actin polymerizations and depolymerisation, protein complex assembly, and regulation of cellular component size (**Table 4B** and lists of DEG (G1vsG2)TN and (G2vsG3)TN in excel file **3_DEGlists_scRNAseq**). This suggests that the blue group of clustered cells, which is the most transcriptionally different to the rest (**Fig. 12**), might already be expressing the intracellular machinery necessary to acquire the dendritic cell morphology and the antigen presenting function characteristic of mature cDC. The

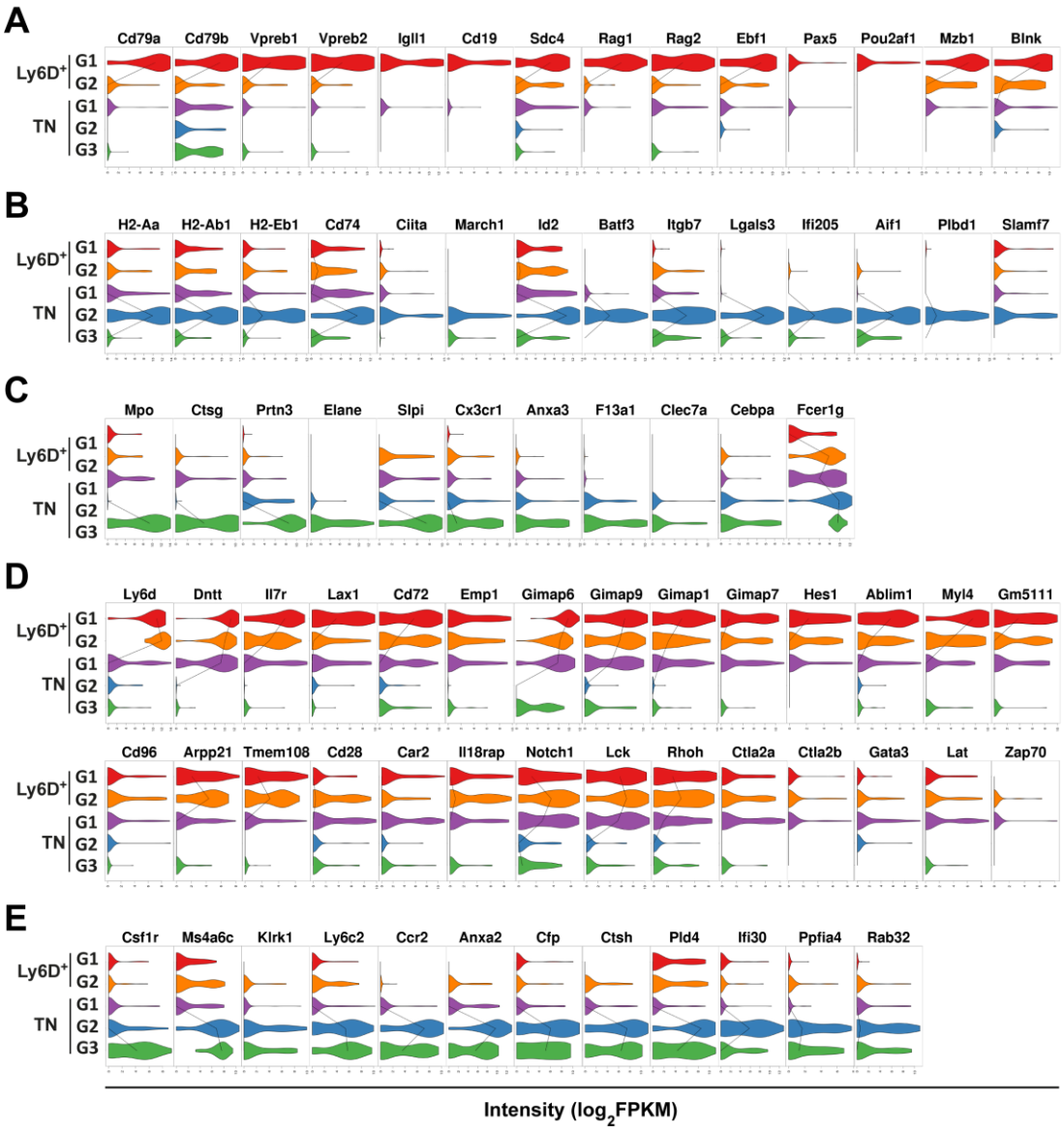


Figure 13. Distinct genetic signatures of the Ly6D$^+$ and TN subgroups. Violin plots with up-regulated or exclusively expressed genes in G1 Ly6D$^+$ **(A),** G2 TN **(B),** G3 TN **(C),** (G1, G2) Ly6D$^+$ and G1 TN **(D),** or G2 and G3 TN **(E)** subgroups**.** The median expression level is shown with a line when more than 50% of the cells express the indicated gene. G1 Ly6D$^+$ (n=56), G2 Ly6D$^+$ (n=82), G1 TN (n=85), G2 TN (n=52), G3 TN (n=56).

G3 TN green subgroup mostly up-regulates myeloid related genes (**Fig. 13C**) involved in innate biological processes (**Table 4C**) characteristic of the entire TN population (**Fig. 9A** lower panel), thus suggesting that this might be the fraction of TN largely containing the observed myeloid potential. However, although the G3 TN subset is myeloid biased, its genetic signature is not resolved to any specific myelomonocytic cell. Some genes such as myeloperoxidase (*Mpo*) are characteristic of both monocyte/macrophages and granulocytes while others are exclusive to granulocytes (*Elane*, expressed in neutrophils) or monocyte/macrophages (chemokine *Cx3cr1*).

Interestingly, our screening did not find genes exclusively expressed in G2 Ly6D$^+$ orange cells or G1 TN purple cells. Instead, we observed that these two subsets follow the same pattern of expression, meaning that when a gene is expressed in one group it is also expressed in the other one (**Fig. 13D**) or it is down-

A

G1 Ly6D$^+$ (compared with G2 Ly6D$^+$)

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0050853~B cell receptor signaling pathway	3	3.16E-04	107.84
GOTERM_BP_FAT	GO:0050851~antigen receptor-mediated signaling pathway	3	0.00213109	41.94
GOTERM_BP_FAT	GO:0042113~B cell activation	3	0.00967014	19.36
GOTERM_BP_FAT	GO:0002253~activation of immune response	3	0.01166008	17.56
GOTERM_BP_FAT	GO:0042100~B cell proliferation	2	0.02273043	83.88
GOTERM_BP_FAT	GO:0042592~homeostatic process	5	0.02380206	4.31
GOTERM_BP_FAT	GO:0050778~positive regulation of immune response	3	0.02762708	11.10
GOTERM_BP_FAT	GO:0045449~regulation of transcription	9	0.05079453	2.03
GOTERM_BP_FAT	GO:0007169~transmembrane receptor protein tyrosine kinase signaling pathway	3	0.05171191	7.86
GOTERM_BP_FAT	GO:0030183~B cell differentiation	2	0.08446759	21.88

B

G2 TN (compared with G1 and G3 TN)

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0002495~antigen processing and presentation of peptide antigen via MHC class II	4	8.59E-05	44.70
GOTERM_BP_FAT	GO:0002504~antigen processing and presentation of polysaccharide antigen via MHC class II	4	1.47E-04	37.64
GOTERM_BP_FAT	GO:0030036~actin cytoskeleton organization	6	0.00216261	6.50
GOTERM_BP_FAT	GO:0007159~leukocyte adhesion	3	0.00262056	38.31
GOTERM_BP_FAT	GO:0008064~regulation of actin polymerization or depolymerization	4	0.00262785	14.30
GOTERM_BP_FAT	GO:0030029~actin filament-based process	6	0.00286076	6.10
GOTERM_BP_FAT	GO:0032271~regulation of protein polymerization	4	0.00381628	12.55
GOTERM_BP_FAT	GO:0043254~regulation of protein complex assembly	4	0.00552056	11.00
GOTERM_BP_FAT	GO:0032535~regulation of cellular component size	5	0.01203336	5.55
GOTERM_BP_FAT	GO:0022610~biological adhesion	9	0.01206098	2.86
GOTERM_BP_FAT	GO:0019882~antigen processing and presentation	4	0.01228952	8.22
GOTERM_BP_FAT	GO:0002573~myeloid leukocyte differentiation	3	0.01504575	15.78

C

G3 TN (compared with G1 TN)

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0001878~response to yeast	3	1.46E-04	154.41
GOTERM_BP_FAT	GO:0009620~response to fungus	3	0.00129085	54.50
GOTERM_BP_FAT	GO:0045807~positive regulation of endocytosis	3	0.00514635	27.25
GOTERM_BP_FAT	GO:0051050~positive regulation of transport	4	0.00833016	9.36
GOTERM_BP_FAT	GO:0006909~phagocytosis	3	0.01046953	18.91
GOTERM_BP_FAT	GO:0006955~immune response	6	0.01594843	3.93
GOTERM_BP_FAT	GO:0060627~regulation of vesicle-mediated transport	3	0.02831563	11.16
GOTERM_BP_FAT	GO:0016337~cell-cell adhesion	4	0.03837022	5.23
GOTERM_BP_FAT	GO:0032760~positive regulation of tumor necrosis factor production	2	0.04645453	41.18
GOTERM_BP_FAT	GO:0006911~phagocytosis, engulfment	2	0.04645453	41.18
GOTERM_BP_FAT	GO:0006952~defense response	5	0.05239538	3.45

D

G1 TN (compared with G2 and G3 TN)

Category	Term	Count	PValue	Fold Enrichment
GOTERM_BP_FAT	GO:0042110~T cell activation	17	1.33E-08	6.18
GOTERM_BP_FAT	GO:0002520~immune system development	26	3.38E-08	3.72
GOTERM_BP_FAT	GO:0048534~hemopoietic or lymphoid organ development	24	2.28E-07	3.60
GOTERM_BP_FAT	GO:0030217~T cell differentiation	12	1.57E-06	6.66
GOTERM_BP_FAT	GO:0042113~B cell activation	12	2.04E-06	6.49
GOTERM_BP_FAT	GO:0046651~lymphocyte proliferation	8	5.10E-05	8.04
GOTERM_BP_FAT	GO:0042100~B cell proliferation	5	1.30E-04	17.58
GOTERM_BP_FAT	GO:0050867~positive regulation of cell activation	11	1.35E-04	4.60
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	13	1.57E-04	3.81
GOTERM_BP_FAT	GO:0050851~antigen receptor-mediated signaling pathway	7	1.78E-04	8.21
GOTERM_BP_FAT	GO:0033151~V(D)J recombination	8	9.07E-04	5.12
GOTERM_BP_FAT	GO:0050856~regulation of T cell receptor signaling pathway	7	0.00135294	5.68
GOTERM_BP_FAT	GO:0030183~B cell differentiation	4	0.00391788	12.06
GOTERM_BP_FAT	GO:0050864~regulation of B cell activation	22	0.00902477	1.83
GOTERM_BP_FAT	GO:0045165~cell fate commitment	4	0.02272059	6.49

Table 4. Selection of enriched Biological Processes (eBP) in up-regulated genes of the G1 Ly6D$^+$ **(A)**, G2 TN **(B)**, G3 TN **(C)**, G1 TN **(D)** groups of cells compared with the subgroups indicated in brackets. Complete list in excel file **(4_eBP_scRNAseq_subgroups)**.

regulated in both (**Fig. 13E**). Moreover, their pattern of expression is qualitatively linked to that of the red subset. Quantitatively, some genes are more highly expressed in the G1 Ly6D$^+$ cells (**Fig. 13D, upper panels**) whereas others in the G2 Ly6D$^+$ and G1 TN cells (**Fig. 13D, lower panels**). Among the latter ones, there are T-cell related genes such as *Notch1*, *Lck*, *Rhoh, Ctla2a*, *Ctla2b, Gata3*, *Lat* or *Zap70* (**Fig. 13D, lower-left panels**), thereby indicating that the orange and purple cells might retain T-cell developmental potential. As a conclusion, these two groups have a lymphoid genetic profile that is not resolved into any particular lineage, presenting B and T enriched biological processes (**Table 4D**) and co-expressing B-, T- and lymphoid-genes (*Il7r*, *Dntt* or *Lax1,* **Fig. 13D** upper left). Finally, there are genes expressed in both green and blue cells (**Fig. 13E**). These are mostly myeloid related genes (*Csfr1*, *Ccr2*, *Ifi30* or *Ctsh*) because although the blue cells are some way along the cDC lineage, as immature DC they have myeloid properties such as phagocytosis.

In summary, the single-cell transcriptomic analysis of the Ly6D$^+$ and TN EPLM subpopulations reveals that: i) the clustered groups of cells have distinct genetic signatures (summarized in **Fig. 14A**) with a remarkable lympho-myeloid separation and ii) the degree of heterogeneity in the entire Ly6D$^+$ and TN populations is reflected into their subgroups' genetic profiles. Thus, both Ly6D$^+$ subsets present a lymphoid genetic profile (with the G1 Ly6D$^+$ cell in a more advanced developmental stage towards the B-cell lineage), whereas those of the TN present signatures to both the lymphoid (G1 TN) or myeloid (G2 and G3 TN) lineages, including some with a cDC lineage profile (G3 TN).

6.7 Expression of lymphoid and myeloid genes is mutually exclusive in single EPLM cells

We have observed that the same group of cells can co-express genes of different lymphoid or myeloid lineages. For instance, the orange G2 Ly6D$^+$ cells express both the T-cell master regulator *Notch1* and the early B-cell transcription factor *Ebf1* (**Fig. 13A,D**), whereas the green G3 TN cells express both the neutrophil marker *Elane* and the macrophage colony-stimulating factor receptor *Csf1r* (**Fig. 13C,D**). In order to elucidate if the previous expression patterns also occur at the

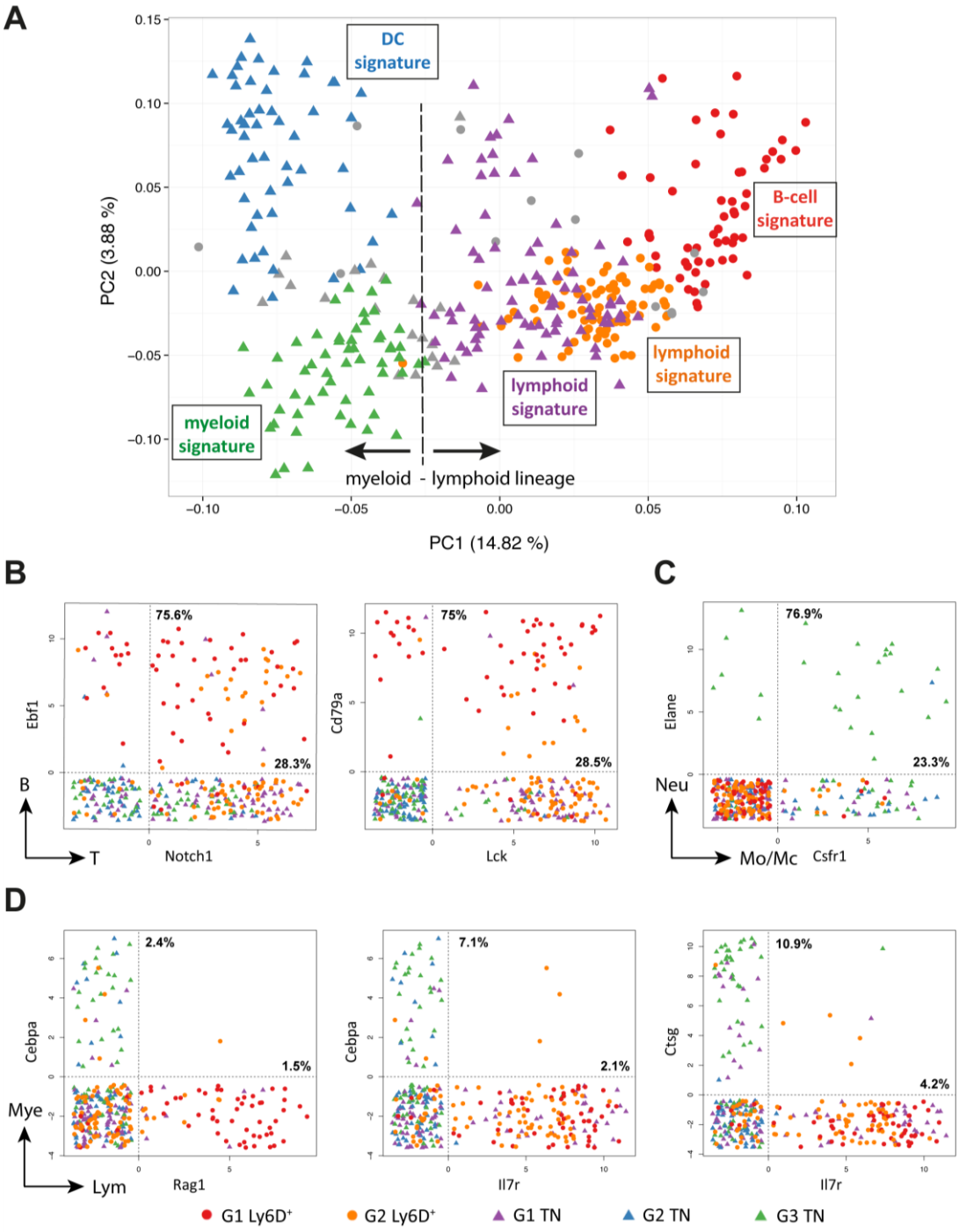


Figure 14. Mixed- and opposed-lineage states at the single-cell level. (A) Same PCA plot as in Fig. 12A summarizing the genetic signatures of the Ly6D$^+$ and TN subgroups revealed by our bioinformatic analysis. **(B-D)** Scatter plots showing the expression levels in $\log_2$FPKM of selected B and T **(B)**, Neutrophil (Neu) and Monocyte/Macrophages (Mo/Mc) **(C)** or Myeloid (Mye) and Lymphoid (Lym) **(D)** lineage marker pairs in the Ly6D$^+$ and TN subpopulations. Bottom legend indicates the colour and shape corresponding to each subgroup. Dotted vertical and horizontal lines delimit when the transcript of the indicated gene is detected (> 0). Percentages within the double-positive area of the plot indicate the fraction of cells co-expressing both genes to the number of cells expressing only one gene (top: gene on vertical axis; bottom: gene on horizontal axis).

single-cell level, we plotted the expression levels of representative pairs of transcripts, each characteristic of different lineages. We observed that a large

proportion of the *Ebf1*$^{+}$ cells also expressed *Notch1* (75,6%, **Fig. 14B** left). The co-expression level was also high when comparing the immunoglobulin α chain of the pre-BCR complex (*Igα* or *Cd79a*) with the T-cell tyrosine kinase *Lck* (75% of the *CD79a*$^{+}$ and 28.5% of the *LcK*$^{+}$ cells, **Fig. 14B** right) Moreover, when examining neutrophil-monocyte/macrophage lineages, a high proportion of single-cells co-expressed *Elane* and *Csfr1* (76.9% of the *Elane*$^{+}$ and 23.3% of the *Csfr1*$^{+}$ cells, **Fig. 14C**). Therefore, the EPLM progenitor population contains single cells with mixed-lineage states within the lymphoid (B and T) and myeloid (granulocyte and monocyte/macrophage) lineages.

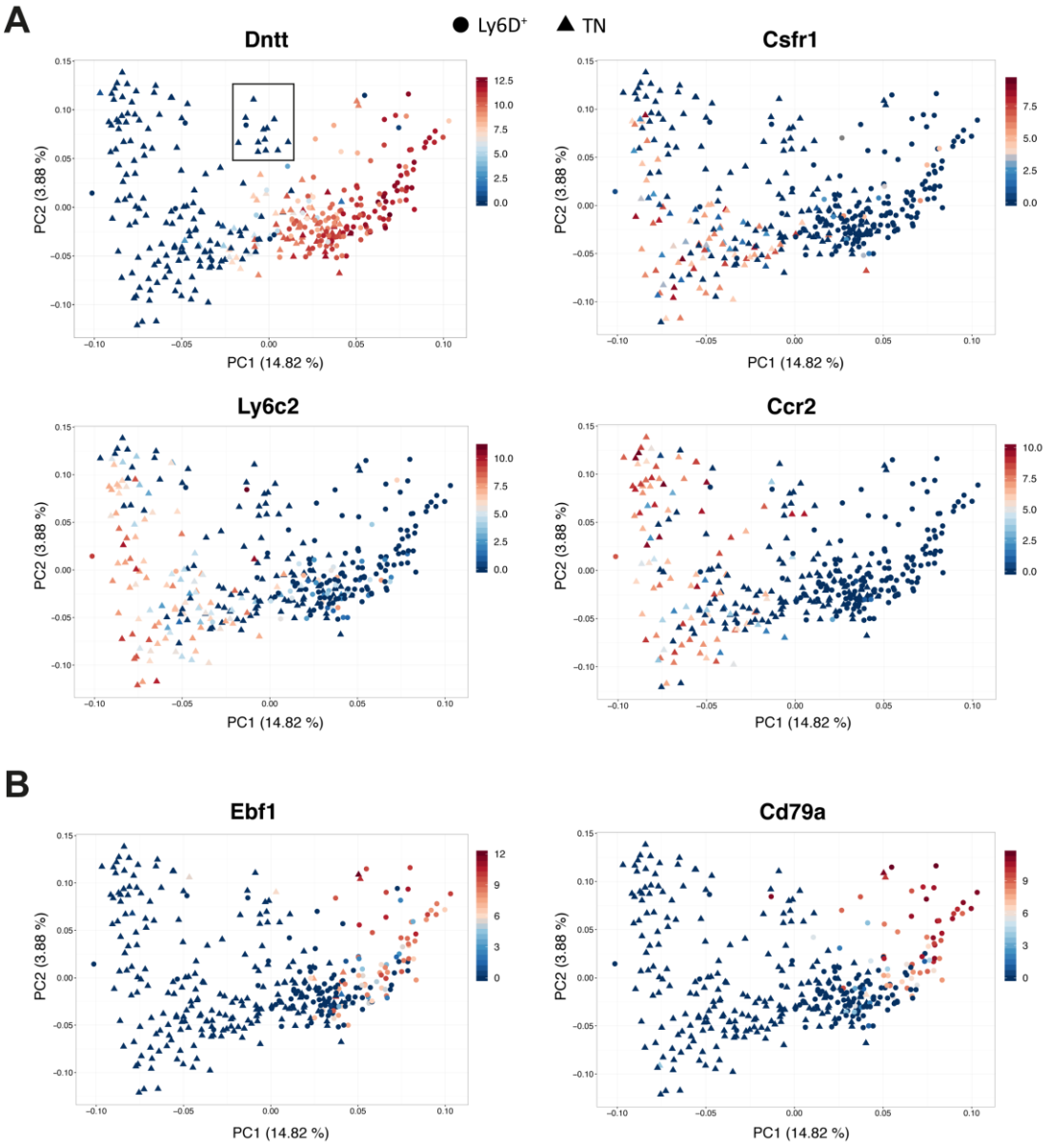


Figure 15. Marker candidates to isolate either the G1 TN (A) or the G1 Ly6D^{+} group of cells (B). PCA plots generated as in Fig. 12A. Colour represents the expression level in log_2FPKM for each indicated gene per cell. Shape indicates the cell type: circles Ly6D^{+}, triangles TN.

In contrast, we detected distinct lymphoid and myeloid specification for the EPLM subgroups, with the cells on the left part of the PCA plot myeloid primed (G2 and G3 TN), whereas those in the centre (G1 TN and G2 Ly6D$^+$) and on the right (G1 Ly6D$^+$), lymphoid primed (**Fig. 14A**). This marked lympho-myeloid separation was confirmed at the single-cell level since we did not encounter significant co-expression of early myeloid (*Cebpa* and *Ctsg*) and lymphoid (*Rag1* and *Il7r*) specification genes (**Fig. 14D**). The mutual exclusive expression of lymphoid and myeloid genes indicates that the EPLM might be composed of a mixture of cells with either lymphoid or myeloid priming and, that the multilineage developmental potential observed for the TN EPLM subpopulation is possibly not contained in the same single cell.

6.8 Selective markers for each subgroup. TdT the best candidate to separate cells with lymphoid from those with myeloid genetic profiles

After identifying selectively expressed markers for each of the clustered groups of cells (**Fig. 13**), we selected the best candidates in order to isolate the subgroups of interest and validate their genetic signatures with functional assays. We first isolated the G1 TN purple cells since we hypothesized that this TN fraction could be the precursor of the Ly6D$^+$ EPLM subpopulation. For that, we performed and extensive screening using the PCA plot as a powerful way of interrogating the gene expression of any gene across the dataset and the best candidate was the *Dntt* gene (**Fig. 15A** upper left). In the PCA plot with cells coloured according to the *Dntt* transcript level (expressed in $\log_2$FPKM), we observe that all G1 and G2 Ly6D$^+$ cells (circles) are *Dntt* positive as well as the vast majority of G1 TN cells (purple in **Fig. 13D** upper left panel and **Fig. 14A**) with the exception of the small upper cluster of cells (box in the PCA plot in **Fig. 15A)**. The *Dntt* gene encodes the Terminal deoxynucleotidyl Transferase (TdT) enzyme, a non-template polymerase that catalyzes the addition of random nucleotides at the junction of rearranged genes of both B- and T-cell receptors, thereby ensuring a highly diverse B- and T-cell receptor repertoire.

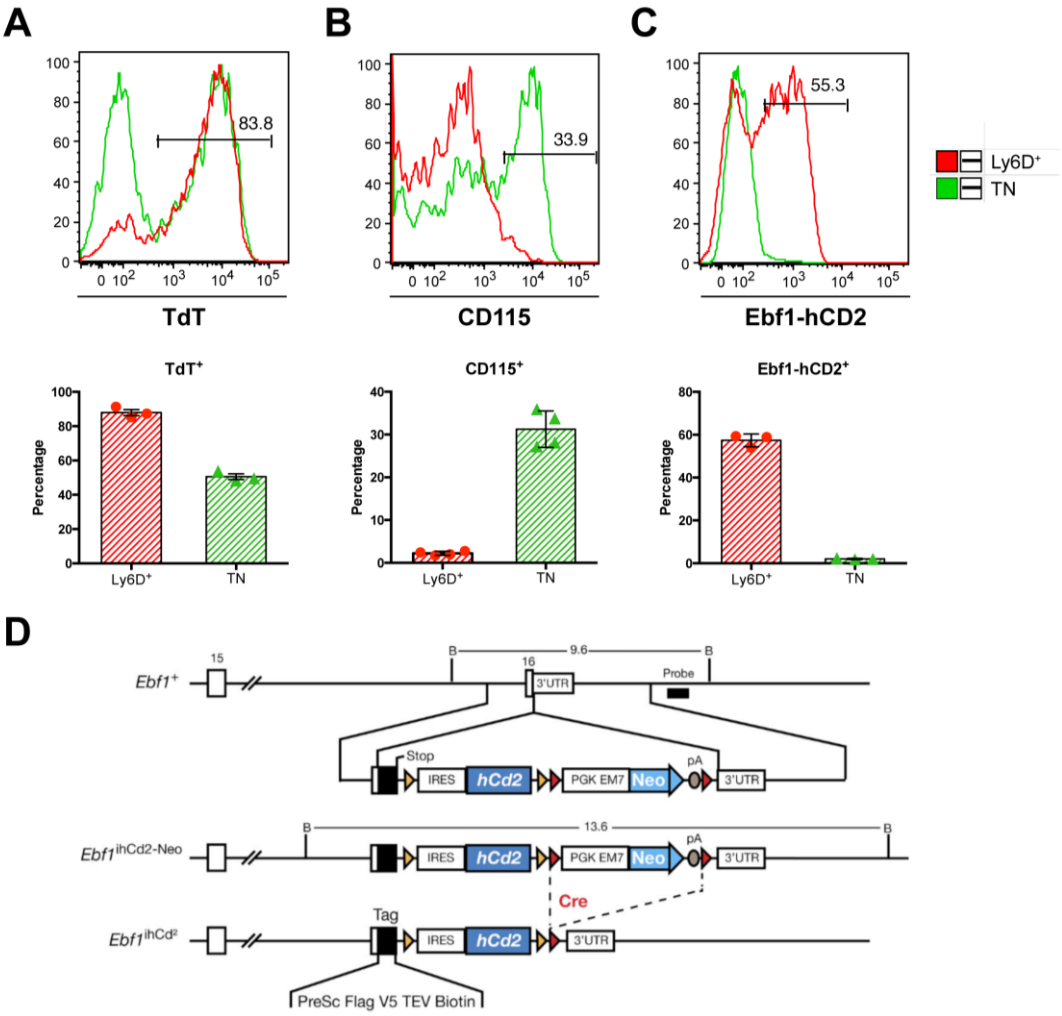


Figure 16. Expression of TdT, CD115 and Ebf1 within the Ly6D$^+$ and TN EPLM subpopulations. (A-C, upper panels) Representative FACS plot showing expression of TdT (A), CD115 (B) or Ebf1 (C) protein within the Ly6D$^+$ and TN EPLM of *Flt3L*tg **(A,B)** or *Flt3Ltg*-Ebf1$^{ihCd2/+}$ **(C)**. (**A-C**, lower panels) Percentages of TdT **(A)**, CD115 **(B)** or Ebf1 **(C)** expressing Ly6D$^+$ and TN EPLM from *Flt3Ltg* (n=3, **A**; n=4, **B**) or *Flt3Ltg* x Ebf1$^{ihCd2/+}$ (n=3, **C**). Shown as mean±SEM. **(D)** Construct to generate the Ebf1$^{ihCd2/+}$ reporter mice as reported in **[112].**

We checked the TdT protein expression by intracellular staining of the EPLM progenitor population and confirmed by flow cytometry that TdT can split the TN (B220$^+$ CD117^{int} CD19$^-$ NK1.1$^-$ Ly6D$^-$ SiglecH$^-$ CD11c$^-$) cells into two almost equal fractions (50.5% TdT$^+$) and that it is expressed by nearly all (87.8%) Ly6D$^+$ EPLM (**Fig. 16A**). Therefore, TdT is potentially an excellent marker to isolate the TN TdT$^+$ cells, equivalent to the G2 TN cluster identified in our bioinformatics analysis, and to test if their developmental potential matches that of their observed lymphoid genetic signature. Unfortunately, TdT is not a cell-surface marker and cannot be directly used to sort TdT positive cells by FACS. Moreover, there is no TdT reporter mouse available. For that reason, our laboratory is in the process of generating a TdT

reporter mouse, which will shed light to this and other ongoing projects in the near future.

As an alternative to TdT, we screened for cell surface candidates that would enable us to enrich for the G1 TN fraction of cells. The most promising were those encoded by *Csf1r*, *Ly6c2* and *Ccr2* genes (**Fig. 15A**), whose expression is mostly overlapping with the blue G2 and green G3 TN cells **(Fig. 13E** and **Fig. 14A)**. Moreover, as shown in the PCA plots (**Fig. 15A**) they are not expressed by exactly the same single cells, suggesting that they could be used in combination in order to exclude a major fraction of cells belonging to the G2 and G3 TN clusters. We attempted several protein surface stainings of the EPLM with individual or combinations of CD115 (or CSF-R, encoded by *Csf1r* gene), Ly6C2 and/or Ccr2 markers. However, we only detected significant expression of the CD115 cell surface protein, with about one third of the TN EPLM subpopulation being $CD115^+$ (**Fig. 16B**).

We also wanted to isolate the G1 $Ly6D^+$ red cells in order to validate their B-cell specification. For that, the two best candidates were the B-cell transcription factor *Ebf1* and the BCR complex-associated protein α chain *Cd79a* (**Fig. 15B and Fig. 13A**). We had access to the *Ebf1*$^{ihCd2/+}$ mice provided by Prof. Meinrad Busslinger and co-workers [111]. These mice had been generated by inserting an internal ribosome entry sequence (IRES)-*hCd2* (ihCd2) reporter gene into the 3' untranslated region of the *Ebf1* gene (**Fig. 16D**). We next crossed the *Ebf1*$^{ihCd2/+}$ reporter to the *Flt3L*tg mice, thereby generating the *Ebf1*$^{ihCd2/+}$*-Flt3Ltg* mice. EPLM staining of the latter mice revealed that on average 57.5% of $Ly6D^+$ cells express hCD2 and therefore Ebf1 (**Fig. 16C**).

6.9 Excluding $CD115^+$ cells enriches for TN cells with lymphoid profile

In order to assess if the detected fraction of $CD115^-$ TN cells possesses the lymphoid genetic signature observed in the single-cell RNA sequencing experiment, we performed several assays. First, we examined the B-cell, T-cell and myeloid precursor frequencies of the two TN ($CD115^+$ and $CD115^-$) fractions. For that, we sorted them from *Flt3L*tg mice and plated graded numbers of cells with the appropriate cocktail of stromal cells and cytokines (same experimental procedure

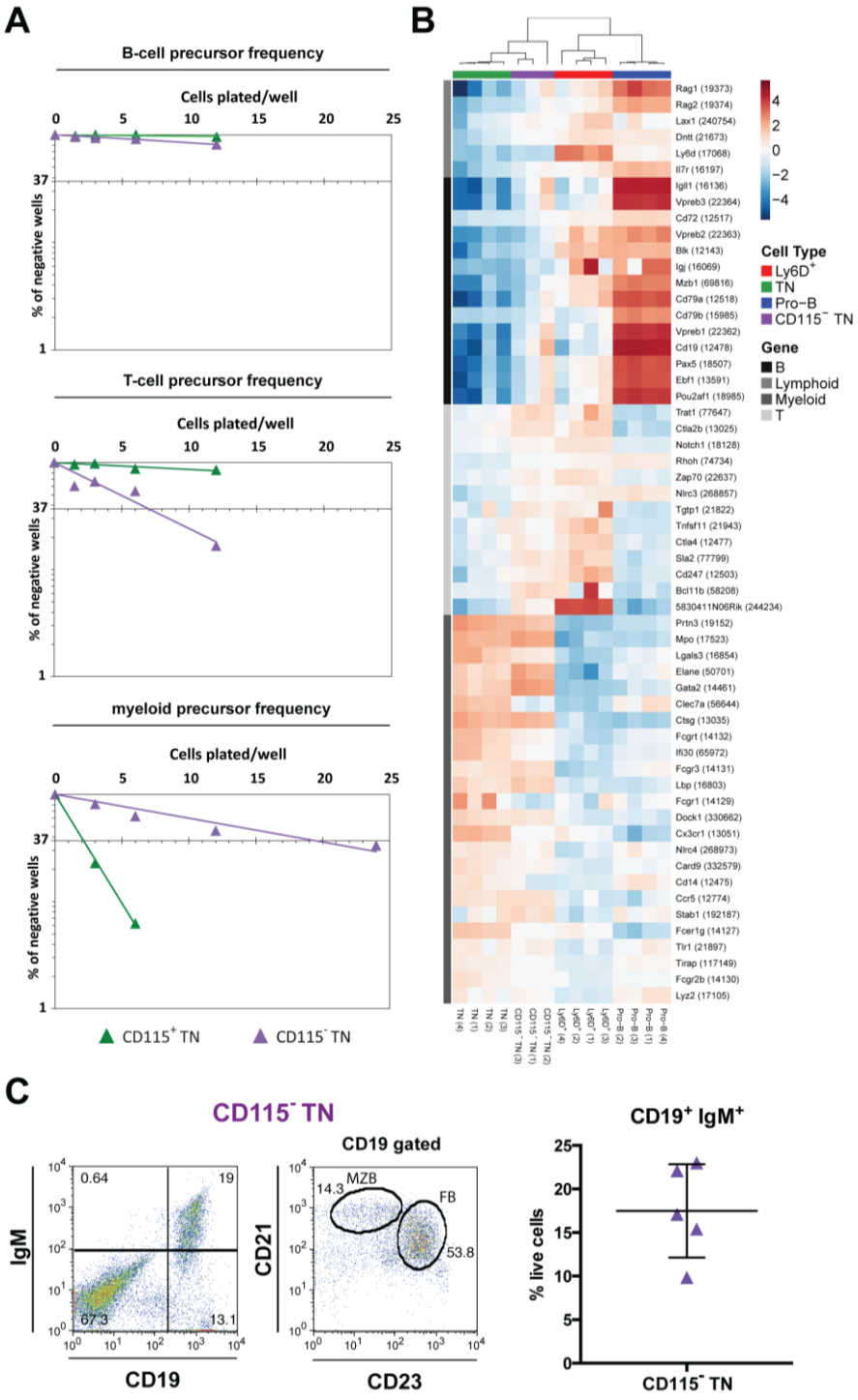


Figure 17. **CD15$^-$ TN fraction contains the lymphoid profile. (A)** Limiting dilution analysis of CD115$^+$ TN and CD115$^-$ TN for B-cell, T-cell or myeloid developmental potentials. EPLM subpopulations were sorted from *Flt3L*tg (2 pooled mice) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7, OP9-DL1 stromal cells in the presence of IL-7, or ST2 stromal cells. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1 and ST2 cell cultures), B-cell, T-cell or myeloid clones were scored using an inverted microscope. One representative experiment is shown out of 2. **(B)** Heatmap of expression of lineage-specific genes as in Fig. 9B. The colour gradient illustrates the centred gene expression level (log_2CPM, mean=0 across all samples and per gene). **(C)** Reconstitution of B-cell compartment in sub-lethally irradiated B6 *Rag2*-deficient mice with 5x10^3 CD115$^-$ TN cells from WT mice. Left panel: CD19 versus IgM expression on spleen cells 3 weeks after transfer. Also CD21 and CD23 expression on gated CD19$^+$ spleen cells are shown (middle panel). Right panel: Quantification of CD19$^+$ IgM$^+$ populations presented as frequency of live cells. CD115$^-$ TN (n=5). Shown as mean±SEM.

explained in results section 2 and **Fig. 5**). The results indicated that the CD115$^-$ TN fraction had lymphoid potential, poor B-cell precursor frequency (< 1 in 50) and strong T-cell precursor frequency (1 in 7), whereas the CD115$^+$ TN did not (**Fig. 17A** upper and middle panel). Moreover, the latter population possessed considerably higher myeloid developmental potential (1 in 2 for CD115$^+$ TN versus 1 in 20 for CD115$^-$ TN) (**Fig. 17A** lower panel). Although low, the myeloid potential observed in the CD115$^-$ TN could be explained because this fraction contains the G1 TN cells as well as others of the G2 and G3 clustered groups (**Fig. 15A** upper-right panel) that showed a myeloid genetic signature. We believe that using TdT as a marker we would be able to discriminate the cells with myeloid developmental potential.

The observed lymphoid developmental potential of the CD115$^-$ TN cells would appear to be in line with their genetic signature, as revealed by bulk RNA sequencing. Briefly, we compared the previously sequenced Ly6D$^+$, TN and Pro-B populations (**Fig. 9**) with the newly identified CD115$^-$ TN fraction. In **Fig. 17B**, a heatmap generated with the same genes as was shown in **Fig. 9B**, indicates that the CD115$^-$ TN subset has a gene expression pattern more similar to the Ly6D$^+$ than the entire TN population has to the Ly6D$^+$ cells, especially for T-cell related genes. Finally, although the CD115$^-$ TN subset had a poor B-cell precursor frequency *in vitro*, they were able to reconstitute the splenic B-cell compartments of *Rag2*-deficient mice (**Fig. 17C**). This finding shows their *in vivo* B-cell developmental potential and suggests that there might be a factor that is necessary to support B lymphopoiesis missing in the *in vitro* cultures.

Collectively, these results indicate that the CD115 marker identified by single-cell RNA sequencing is able, as predicted by the genetic profiles of the clustered groups of cells (**Fig. 14A**), to exclude cells with myeloid precursor frequency and thereby enrich for those having lymphoid profiles and being the precursor of the Ly6D$^+$ EPLM subpopulation.

6.10 Ebf1 enriches for Ly6D$^+$ cells with B-cell potential

In order to elucidate if the newly identified Ebf1$^+$ Ly6D$^+$ subset is B-cell primed as predicted by the single-cell RNA sequencing experiment, we assessed their

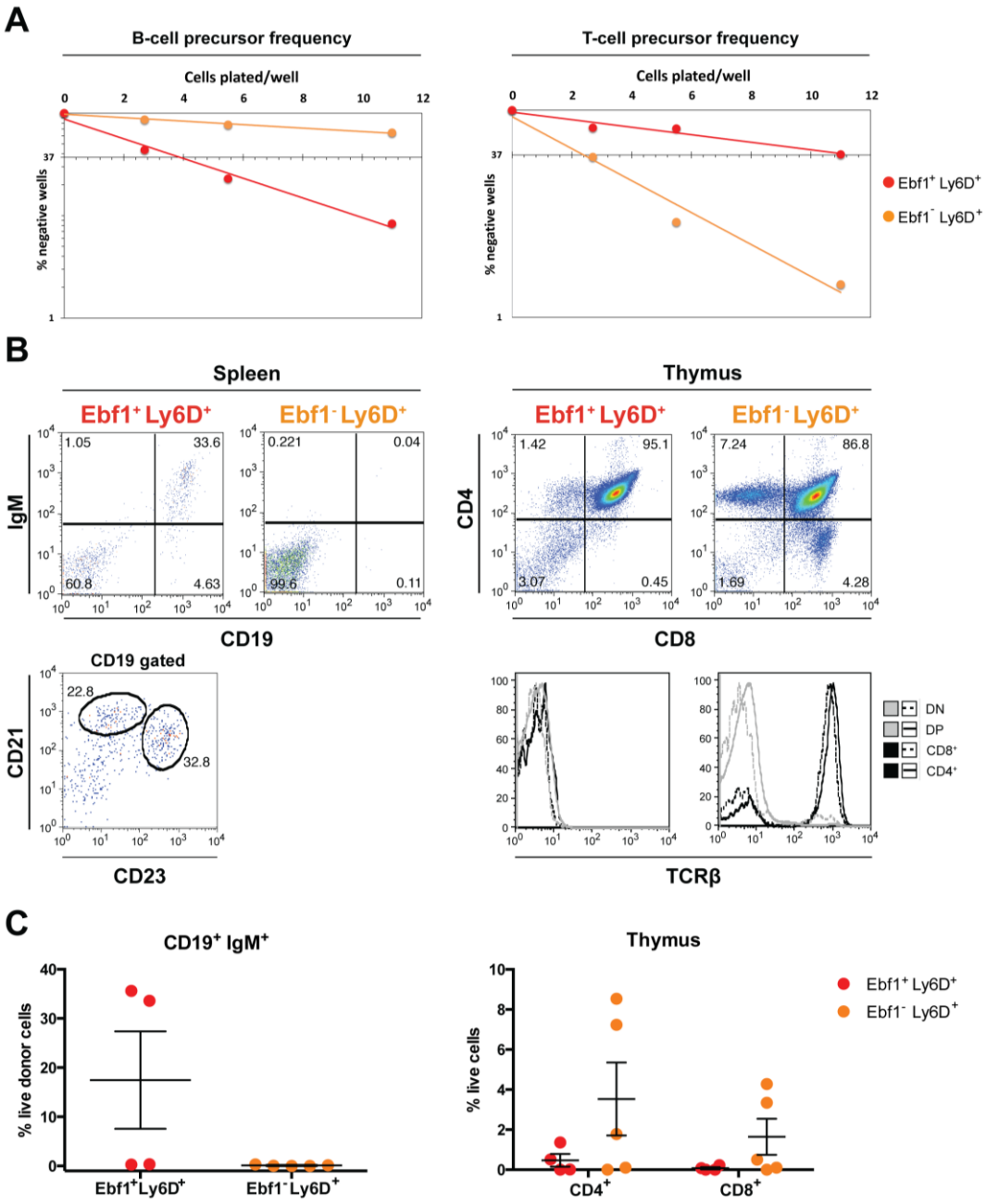


Figure 18. Ebf1$^+$ Ly6D$^+$ fraction is enriched in B-cell whereas Ebf1$^-$ Ly6D$^+$ fraction in T-cell developmental potential. (A) Limiting dilution analysis of Ebf1$^+$ Ly6D$^+$ and Ebf1$^-$ Ly6D$^+$ for B-cell and T-cell developmental potentials. EPLM subpopulations were sorted from *Flt3L*tg-Ebf1$^{ihCd2/+}$ (2-pooled mice) and plated at the indicated concentrations on either OP9 stromal cells together with IL-7 or OP9-DL1 stromal cells in the presence of IL-7. After 10 days (for OP9 cell cultures) or 15 days (for OP9-DL1), B-cell or T-cell clones were scored using an inverted microscope. One representative experiment is shown out of 2. **(B)** Reconstitution of B (left) and T (right) cell compartment in sub-lethally irradiated B6 *Rag2*-deficient mice with 5x10^3 Ly6D$^+$ Ebf1$^+$ or Ly6D$^+$ Ebf1$^-$ cells from *Flt3L*tg-Ebf1$^{ihCd2/+}$ mice. Left panels: CD19 versus IgM expression on spleen cells 3 weeks after transfer. Also CD21 and CD23 expression on gated CD19$^+$ spleen cells are shown. Right panels: CD4 and CD8 expression on thymocytes 3 weeks after transfer. Also TCRβ expression on DN, DP, CD8$^+$ and CD4$^+$ gated thymocytes is shown. **(C)** Quantification of CD19$^+$ IgM$^+$ (left), CD4$^+$ and CD8$^+$ (right) populations presented as frequency of live donor cells (left) or frequency of live cells (right). Ebf1$^+$ Ly6D$^+$ (n=4), Ebf1$^-$ Ly6D$^+$ (n=5). Shown as mean±SEM.

in vitro and *in vivo* developmental potentials. Since the entire Ly6D$^+$ EPLM subpopulation is already lymphoid restricted, we focused our analysis on examining

the B-cell and T-cell precursor frequencies of the two (Ebf1$^+$ and Ebf1$^-$) Ly6D$^+$ fractions by limiting dilution (**Fig. 18A**). The results indicate that although both subsets have *in vitro* B and T cell developmental potential, the Ebf1$^+$ Ly6D$^+$ fraction has a higher B-cell precursor frequency (1 in 5 for Ebf1$^+$ Ly6D$^+$ versus 1 in 21 for Ebf1$^-$ Ly6D$^+$), whereas the Ebf1$^-$ Ly6D$^+$ fraction has a higher T-cell precursor frequency (almost 1 in 2 for Ebf1$^-$ Ly6D$^+$ versus 1 in 11 for Ebf1$^+$ Ly6D$^+$) (**Fig. 18A**). These results were confirmed by their differential abilities to reconstitute mouse B- and T-cell compartments *in vivo* (**Fig. 18B,C**). Whereas the Ebf1$^+$ Ly6D$^+$ subset had all the B cell reconstitution potential (2/4 vs 0/5), the Ebf1$^-$ Ly6D$^+$ were the only cells able to reconstitute the thymus with TCRβ$^+$ T cells (2/5 vs 0/4) (**Fig. 18C**). Overall, the genetic profiles of the two Ly6D$^+$ clusters (red and orange in **Fig. 14A**) predicted by single-cell RNA sequencing, and exemplified with the Ebf1 marker, are reflected in their developmental potentials.

7. Discussion

The identification of two complementary clonogenic hematopoietic progenitors, CLP in 1997 [27] and CMP in 2000 [28], with the capacity to differentiate into all cells of the lymphoid and myeloid lineages respectively, resulted in the so-called "classical" model of haematopoiesis found in most current text books. This classical model reflects a hierarchical decision-making process whereby early multipotent progenitors make an early and irrevocable decision to differentiate towards either the lymphoid or the myeloid lineages. Since then, the classical model of haematopoiesis is in constant revision as a consequence of finding numerous additional progenitor cell types and alternative routes of differentiation. For instance, the identification of the LMPP [35] suggests an earlier branching of the megakaryocyte/erythrocyte pathway to the lympho-myeloid lineages, before the branching of the two latter lineages, and is supported by several later studies [70, 135]. Moreover, the irreversible and strict compartmentalization of haematopoietic progenitor cells was first challenged by the finding that in WT mice, committed precursor B cells regain multipotentiality upon conditional inactivation of the *Pax5* gene [31]. This suggested that there might be a greater degree of developmental plasticity in haematopoiesis than previously thought. Subsequently, our laboratory detected in the BM of wild type mice a cell with similar properties to the *Pax5*$^{-/-}$ Pro-B cells that we called "Early Progenitor with Lymphoid and Myeloid potential" (EPLM) [33]. This was an uncommitted and multipotent haematopoietic B220^{+} CD117^{int} CD19^{-} NK1.1^{-} progenitor with combined lymphoid and myeloid developmental potential, therefore challenging the lympho-myeloid dichotomy supported by the classical model of haematopoiesis.

In the present study, we have interrogated the multipotentiality of the EPLM in detail and assessed its potential heterogeneity. First, we found that EPLM expressed heterogeneous levels of three cell surface markers Ly6D, SiglecH and CD11c, resulting in subdivision of EPLM into at least four subpopulations (**Fig. 4D**). Moreover, this EPLM subdivision was still present in *Flt3L*tg mice, in which the numbers of all EPLM subpopulations were significantly increased almost equally (about two orders of magnitude, **Fig. 4E**), thus making the *Flt3L*tg mouse a good

model to further study EPLM subpopulations. When we assessed the *in vitro* and *in vivo* developmental potential of individual subsets by limiting dilution and reconstitution assays respectively, we observed that EPLM heterogeneity was reflected in different sets of potentials. Briefly, SiglecH$^+$ and CD11c$^+$ subpopulations could not generate lymphoid cells and, as their cell surface marker profile suggests, they could be (at least a fraction) already committed to the pDC and cDC lineages. Ly6D$^+$ cells were lymphoid restricted and TN showed trilineage (B, T and myeloid) developmental potential (**Fig. 5**) although with lower B- and T-cell precursor frequencies compared with their Ly6D$^+$ counterparts (**Fig. 5A,B**). Of note, Ly6D$^+$ and TN cells from *Flt3L*tg mice showed lower B-cell potential but higher T-cell precursor frequencies (**Fig. 5A,B**), as well as increased T-cell reconstitution efficacy in the thymus (**Fig. 6D**). One explanation for this finding could be the lower levels of Pax5 and Ebf1 B-cell transcription factors detected in Ly6D$^+$ cells from *Flt3L*tg mice compared to their WT counterparts (von Muenchow et al., **Appendix paper 1**). However, Flt3L is not an instructive factor for B-cell commitment of these progenitors. Instead, the main effect of Flt3L is the induction of their proliferation (von Muenchow et al., **Appendix paper 1**). Therefore, it could well be that an increased percentage of cycling progenitors results in a decreased fraction initiating the B-cell developmental program, thus explaining the proportional reduction of Ebf1- and Pax5-expressing Ly6D$^+$ cells and the increased ability of these progenitors to give rise to T cells. However, it is important to mention that, in *Flt3L*tg mice, the total number of B-cell progenitors is not affected due to the general expansion of the EPLM compartment (**Fig. 4**). Since Flt3L is reported to be important for pDC development [109, 136, 137] and the EPLM subpopulations might have pDC developmental potential, we hypothesize that pDC progenitors might be expanded in the transgenic mice, thereby providing an alternative explanation for the reduced B-cell potential. Although a clonal assay to assess quantitative pDC precursor frequency does not yet exist, it would be of interest to investigate the pDC developmental potential of EPLM subpopulations from WT and *Flt3L*tg mice.

Another striking observation from the developmental potential assays was that, in spit of the poor *in vitro* capacity of TN cells to generate B cells (**Fig. 5A**), they were nevertheless able to reconstitute the splenic B-cell compartments of all

injected mice (**Fig. 6C**). This result therefore suggests that in *in vitro* cultures, there might be a factor missing that is necessary to support B lymphopoiesis. The opposite was observed for the T-cell potential in that this was more efficient *in vitro* than *in vivo* (**Fig. 5B** and **Fig. 6D**). A key difference in the two experimental systems used to determine T-cell potential is the availability of Notch ligands. Whereas the *in vitro* system generates highly sensitive readouts that take advantage of the T-cell lineage-promoting activity of Notch1 signalling via culturing progenitors on feeder cells expressing Notch ligands [138], the intravenous injection system depends on progenitors travelling to the thymus prior to receiving Nocth ligand signalling upon thymus entry. Experimentally, this limitation could be overcome if progenitors were directly injected intrathymically. However, EPLM are progenitor cells residing in the bone marrow and therefore it is probable that although Ly6D$^+$ cells show robust *in vitro* T-cell potential, their "default" physiological fate, when no other lineage commitment signal, such as Notch ligand, interferes, might be to become a B cell. Finally, we observed similar *in vivo* capacity of Ly6D$^+$ and TN progenitors to reconstitute the splenic B-cell compartment (**Fig. 6C**). To further confirm this, it would be worthwhile injecting limiting numbers of progenitors and monitoring their reconstitution at different time points. Overall, these results and associated discussions raise the fundamental question of which assays truly represent the physiological potential of progenitor cells.

Transcriptomic analysis of the Ly6D$^+$ and TN EPLM subpopulations showed that their distinct developmental potentials were supported by differential gene expression programs. Thus Ly6D$^+$ EPLM had a lymphoid molecular priming and transcriptomic profile closer to committed Pro-B cells, whereas the TN had a more myeloid genetic signature (**Fig. 9** and **Table 3**). Even though TN cells up-regulate myeloid genes involved in innate biological processes, we were able to demonstrate the precursor-product relationship between TN and Ly6D$^+$ EPLM subpopulations, with the TN (or at least a fraction thereof) being the precursor of the Ly6D$^+$ subpopulation, and therefore explaining the slower kinetics and decreased efficiency of the latter progenitor differentiating into CD19$^+$ B cells (**Fig. 13**). Collectively, the previous data reveals that the TN EPLM subpopulation is a lympho-myeloid

progenitor, whereas the Ly6D$^+$ EPLM subpopulation is likely to be the direct precursor of CD19$^+$ committed Pro-B cells.

The emergence of high throughput methods enabling the investigation of whole-genome or whole-transcriptome profiles from single cells, boosts the existing debate regarding the heterogeneity of apparent phenotypically homogeneous progenitor cells that have multiple lineage potentials. To shed light onto this active debate, by performing single-cell RNA sequencing we provide herein a detailed analysis of the heterogeneity of the Ly6D$^+$ and TN EPLM subpopulations. This study captured 365 single-cell gene expression snapshots of the Ly6D$^+$ (152) and TN (213) transcriptional landscapes. Principal component analysis is a good method to visualize and interpret high dimensional data (e.g. 365 single cells x 14814 detected genes). This method maximizes the variance of the features (genes in our case). Therefore, in order to obtain informative components (or dimensions) in the plots, it is common practise to select a subset of the original gene set that contains most of the variation among the samples [53, 139, 140]. One strategy is to perform an unsupervised selection of, for example, the 50% of genes with the highest variance across the analysed dataset, as was done for the visualization of the bulk RNA sequencing experiment. However, to analyse the single-cell data we decided to concentrate on the 1008 genes that we identified as differentially expressed between the Ly6D$^+$ and TN populations. Our reasoning was that this set of genes would provide insights into the intra-population variation. This is because the population approach obscures molecular heterogeneity by averaging gene expression. By analysis of bulk populations, high expression of a particular gene in an individual subset will be interpreted as general expression by all cells. As suggested by their branching structure in the PCA and their low cell-to-cell transcriptome correlation our initial analysis indicated that the TN subset is a more heterogeneous population than their Ly6D$^+$ partners (**Fig. 11**). Moreover, the PAM clustering based on the selected subset of genes, partitioned the Ly6D$^+$ into two major clusters whereas the TN was subdivided into three robust groups. Of note, Ly6D$^+$ and TN cells had not been completely segregated in the PCA (**Fig. 11B**), thereby suggesting that the overlapping cells might be in a related cellular state. This was confirmed by the analysis of the two clustered groups corresponding to the convergent area, namely

G2 Ly6D$^+$ and G1 TN (**Fig. 12A**), revealing that they have the highest transcriptome correlation among the subgroups (0.721, **Fig. 12B**) and the lowest number of DEG (25, **Fig. 12C**). Therefore, the G1 TN is likely to be the fraction that we previously observed experimentally to be the precursor of Ly6D$^+$ cells. However, our latter *in silico* analysis performed at the whole-transcriptome scale of the G2 Ly6D$^+$ and G1 TN subgroups, which are sorted as two phenotypically distinct populations, suggests that they should be considered as one subset. This finding highlights the limitations of accurately defining complete cellular identities by relying on expression of few cell surface markers, the so-called "top-down" approach [141]. This notion has been manifested in other studies such as when Paul et al. suggested that the standard gating for sorting MEP might be better termed erythrocyte progenitor (EP) gating [69].

When performing cell clustering, it is of pivotal importance to investigate the nature of the subgroups in order to determine if the cells are clustered as a result of stochastic fluctuations of the transcriptome and certain cellular states (e.g. cell cycle status or cellular stress), or instead, the clustering reflects gene expression variation that might be significant for the biological function being tested [142]. In our case, the hypothesis is to ask whether individual EPLM are multipotent or whether as a population they are composed of a mixture of cells with distinct linage developmental potentials. Therefore, we sought to determine if the clustered groups of cells reflect genetic signatures of progenitors from distinct haematopoietic lineages. For that, we extended our analysis to all detected genes (14,814) across the 365 single-cells and examined the DEG and enriched biological processes defining each subgroup (**Table 4** and supplementary excel file **3_DEGlists_scRNAseq**). With this approach, we were able to unravel marked genetic biases between EPLM indicative of molecular priming towards distinct fates. Whereas the G1 Ly6D$^+$ subgroup showed a strong B-cell genetic signature (red **Fig. 14A**) with robust expression of B-cell related genes characteristic of the Pro-B cell stage (*Cd79a*, *Vpreb* genes, *Igll1*, *Cd19*, *Ebf1* or *Blnk* **Fig. 13A**) and B-cell enriched biological processes **(Table 4A**), the G2 TN subgroup, which has the most distant transcriptome, revealed a cDC genetic signature (blue **Fig. 14A**) with consistent expression of genes (*H2*-genes, *Cd74*, *Ciita*, *Id2* or *Batf3* **Fig. 13B**) and enriched biological processes (**Table 4B**)

either related with antigen processing and presentation or necessary for cDC development. The G3 TN subgroup exhibited a myeloid genetic signature (green **Fig. 14A**), up-regulating myeloid genes (*Mpo*, *Ctsg*, *Prtn3*, *Elane*, *Cx3cr1*, *Cebpa* or *Csfr1* **Fig. 13C,E**) related with innate processes (**Table 4C**). Finally, the G2 Ly6D^{+} / G1 TN subset (orange and purple respectively in **Fig. 14A**) showed a lymphoid genetic signature (**Fig. 13D** and **Table 4D**) with a B-cell specification milder to that of the G1 Ly6D^{+}. Taken together, these results indicate that, as reflected by their central location in the PCA, these cells might me in an intermediate state while exhibiting promiscuous expression of B- and T-cell genes. Therefore, the previous data suggest functional heterogeneity among the EPLM subpopulations and validates single-cell RNA sequencing as a powerful technology to dissect molecular heterogeneity of a previously considered homogeneous population via generating biologically meaningful clusters of cells. In addition, numerous studies have reported the use of this technology with similar purposes [46, 53, 57, 67-69, 139, 143-145]. For instance, Gren et al. reveal transcriptomic variation within the traditionally classified classical, intermediate and non-classical monocytes [67]; Paul et al. report heterogeneity in myeloid progenitors [69]; Kowalczyk et al. find extensive transcriptome variability among HSCs [145]; and Drissen et al. subdivided the pre-GM population into a Gata1^{+} pre-GM fraction generating mast cells, eosinophils, megakaryocytes and erythroid cells and a Gata1^{-} pre-GM subset generating monocytes, neutrophils and lymphocytes [139]. These publications challenge the existence of a clonal myeloid progenitor generating all innate myeloid immune cell types.

We also investigated lineage priming at the single-cell level and found a significant proportion of single cells co-expressing early B- and T-cell (**Fig. 14B**) or granulocyte and monocyte/macrophage (**Fig. 14C**) specification genes. This is consistent with other studies where mixed lymphoid [126, 146] or myeloid [135, 147, 148] lineage patterns of gene expression are reported in single cells. However, whereas heterogeneity is well studied in myeloid progenitors, we are not aware of other reports addressing mixed lymphoid priming at the single-cell level and whole-transcriptome scale, thus highlighting the importance of our study. Strikingly, and in agreement with other reports [144, 149-151], we did not observe single cells with mixed myeloid and lymphoid lineage gene expression profiles (**Fig. 14D**).

Hans-Reimer Rodewald and co-workers made use of an *Il7r* fate mapping mouse line to determine the almost exclusive lymphoid progeny of *Il7r* expressing cells, thus further supporting our findings [149]. Nevertheless, we cannot exclude that, due to the "snapshot" nature of the transcriptomic analysis, as well as the medium throughput of cells analysed and the medium sequencing depth, we are missing a transient and presumably rare intermediate state with promiscuous lympho-myeloid gene expression. Previous data suggested that bifurcation of the lymphoid and myeloid molecular priming occurs before the EPLM stage. Therefore, the common or separate origin of the Ly6D$^+$/G1TN (lymphoid primed) versus the G2/G3TN (myeloid primed) EPLM fractions is of interest and requires further investigation. In line with our findings, there is an increasing body of evidence supporting the notion that priming occurs much earlier in development than previously thought. Indeed, expression of lineage-affiliated genes has been reported as early as in the HSC stage, with various analyses indicating biases at the apex of haematopoiesis [135, 144, 152-158], thereby suggesting that the cellular output of the multipotent early progenitors is small [135].

One key aspect of single-cell RNA sequencing studies is the identification of markers representative of the identified clusters of cells in order to prospectively isolate them and perform functional assays to validate the predicted genetic signatures. In our study, we found that the reclassification of EPLM subpopulations was exemplified by expression of *Dntt* (encoding TdT), which was expressed in the Ly6D$^+$ as well as the vast majority of G1 TN but not in G2 and G3 TN cells (**Fig. 15A**). We believe that TdT reporter mice (being currently generated in our laboratory), will provide the exciting opportunity not only to isolate the lymphoid-primed EPLM fraction but also to test and demonstrate, as previously suggested [159], if TdT$^+$ CD19$^-$ cells constitute the earliest lymphoid progenitor in adult mice. Moreover, this reporter will be of importance in studies comparing fetal versus adult haematopoietic progenitors and to stablish when TdT is first expressed in ontogeny.

Profiting the power of single-cell RNA sequencing to identify markers in an unbiased manner, we performed an extensive screening and found potential candidates (**Fig. 15A,B**). As a proof-of-concept, we made use of *Ebf1*$^{ihCd2/+}$ reporter mice in order to test the developmental potential of Ebf1$^+$ EPLM (comprising most of

the G1 and some G2 Ly6D$^+$ cells). As expected, the Ebf1$^+$ Ly6D$^+$ fraction exhibited higher B-cell precursor frequency and *in vivo* reconstitution potential than the Ebf1$^-$ Ly6D$^+$ fraction (**Fig. 18**). This result demonstrates that we can find markers from our single-cell transcriptomic analysis that, when tested, successfully predict developmental outcome. Although the G1 Ly6D$^+$ fraction has a strong B-cell genetic specification and developmental potential, the observed marginal *in vitro* T-cell potential suggests that they are still not restricted to the B-cell lineage. This supports the concept that genetic specification precedes commitment [147, 160, 161] and that Ebf1 initiates B-cell specification whereas Pax5 acts later to drive commitment [126, 160]. Of note, in our single-cell analysis, we barely detected *Pax5* expression (n=8), which could be due to the insufficient sequencing depth to detect low expressing transcripts, or the use of *Flt3L*tg mice (in which Pax5 is lower at the EPLM stage, von Muenchow et al. **Appendix paper 1**). However, even though EPLM are sorted as CD19$^-$, we consistently detected *Cd19* transcript (n=19). Therefore, we propose the herein newly identified G1 Ly6D$^+$ subset as the direct precursor of the first B-cell committed stage, namely the CD117$^+$ CD19$^+$ Pro-B cells.

Additionally, we tested the CD115 cell surface marker (encoded by *Csf1r* gene) identified in our single-cell RNA sequencing experiment and expressed by a good proportion of the G2 and G3 TN cells. Interestingly, our molecular and functional analysis revealed that compared with the entire TN population, CD115$^-$ TN had reduced myeloid but enriched lymphoid properties (**Fig. 17**), thus confirming that priming at the molecular level correlates with the immediate differentiation bias of EPLM cells. Therefore, our data support intrinsic differentiation potentials through differential gene expression programs, although we do not exclude the contribution of other factors such as epigenetic regulation or extrinsic cytokine signalling.

Single-cell transcriptome profiling has already enabled advances to be made in a broad variety of fields. However, it implies a subsequent prospective strategy which, although useful for this and other research studies [65, 139], depends on the identification of cell-surface markers or the existence of reporter mice in order to translate the single-cell transcriptome findings. Additional genetic tools and integration of different omics analysis will certainly be developed, thus enhancing the contribution of single-cell analysis technology to the field of cell lineage

determination. Indeed, the first genome and transcriptome sequenced from the same individual cell already exists [162], enabling a direct comparison of genomic variation and transcriptome heterogeneity. The challenge will be to generate methods for RNA and protein measurements without destroying the cell of interest in order to relate gene expression with cell behaviour. These measurements will require transcriptional and translational reporters combined in single-molecule imaging and quantification experiments.

As a summary, in this thesis, by studying the previously reported multipotent and uncommitted B220^{+} CD117^{int} CD19^{-} NK1.1^{-} EPLM progenitor, we have been able to: first divide the EPLM into four subpopulations with distinct lineage biases using the "top down" strategy. This is based on the expression of cell surface markers (Ly6D, SiglecH and CD11c in our case). Subsequently, we further investigated the EPLM heterogeneity using the alternative and complementary "botton up" approach [141], in which high-dimensional molecular profiles measured for each single cell are used to classify distinct states and types. The latter strategy has revealed that: i) Ly6D^{+} and TN EPLM subpopulations have distinct degrees of molecular and functional heterogeneity ii) the Ly6D^{+} subset is composed of two lymphoid specified subgroups with a B-cell priming gradient iii) the TN subset is composed of three groups of cells with lymphoid or myeloid genetic signatures, including some cells with a cDC lineage profile iv) genetic signatures reveal a remarkable lympho-myeloid separation and v) this separation is better exemplified by *Dntt* instead of Ly6D, thereby redefining the initial "top down" classification. Taken together, these results indicate that the EPLM is a good example supporting the finding that previously characterized multipotent progenitor populations are in fact composed of mixtures of cells with differently restricted differentiation capacities. EPLM share their B220, CD117 and CD19 phenotype with *Pax5*$^{-/-}$ Pro-B cell clones that exhibit multilineage differentiation capacities. However, despite this phenotypic similarity, we can conclude from the analysis of the developmental potential of individual subsets that EPLM are in fact distinct from *Pax5*$^{-/-}$ Pro-B cells [30, 31, 33]. More generally, this study illustrates how single-cell expression profiling can be used to identify novel subpopulations and cellular intermediates that can then be prospectively isolated and characterized in functional assays.

Finally, in the context of constant and necessary revisions of the haematopoietic schemes, we argue in favour of a revised model of haematopoiesis in which lineage priming is initiated as early as at the HSC level and progressively increases with development, ultimately giving rise to the differentiated blood cell types. Thus, unlike the classical hierarchical and branching model of haematopoiesis, we support alternative representations like the pair-wise model (**Fig. 2**; [41]) where the route a progenitor cell will take to arrive at a particular destination is more flexible than currently represented. This more flexible model better reflects haematopoietic versatility and can potentially accommodate the increasingly reported alternative differentiation pathways [36, 70, 139, 158]. However, our findings suggest a revision of the CLP/EPLM continuum arch of the pair-wise model. Now we know that the lymphoid priming of the EPLM resides in the TdT^+ fraction (herein identified as the $Ly6D^+$ as well as the G1 TN EPLM subpopulations), and therefore this is the fraction of EPLM that should be placed directly downstream of CLP rather than the entire EPLM population [33]. The location of the other EPLM fractions in the above scheme has yet to be defined and requires further investigation. Moreover, haematopoiesis seems to occur through a process of "graded commitment" [69, 135] and this is consistent with our finding of an increasing degree of B-cell priming within EPLM subsets (G2 $Ly6D^+$ / G1 TN < G1 $Ly6D^+$). How to illustrate the global haematopoietic process of "graded commitment" in the revised models of haematopoiesis remains an exciting challenge for the near future. Ultimately, this thesis makes a significant contribution to the study of the phenotypic and genotypic heterogeneity and lineage priming of progenitors during early stages of haematopoietic development.

8. References

1. Birbrair, A. and P.S. Frenette, *Niche heterogeneity in the bone marrow.* Ann N Y Acad Sci, 2016. **1370**(1): p. 82-96.
2. Yin, T. and L. Li, *The stem cell niches in bone.* J Clin Invest, 2006. **116**(5): p. 1195-201.
3. Chotinantakul, K. and W. Leeanansaksiri, *Hematopoietic stem cell development, niches, and signaling pathways.* Bone Marrow Res, 2012. **2012**: p. 270425.
4. Wilson, A. and A. Trumpp, *Bone-marrow haematopoietic-stem-cell niches.* Nat Rev Immunol, 2006. **6**(2): p. 93-106.
5. Oh, I.H. and K.R. Kwon, *Concise review: multiple niches for hematopoietic stem cell regulations.* Stem Cells, 2010. **28**(7): p. 1243-9.
6. Coskun, S. and K.K. Hirschi, *Establishment and regulation of the HSC niche: Roles of osteoblastic and vascular compartments.* Birth Defects Res C Embryo Today, 2010. **90**(4): p. 229-42.
7. Avecilla, S.T., et al., *Chemokine-mediated interaction of hematopoietic progenitors with the bone marrow vascular niche is required for thrombopoiesis.* Nat Med, 2004. **10**(1): p. 64-71.
8. Brummendorf, T.H., et al., *Asymmetric cell divisions sustain long-term hematopoiesis from single-sorted human fetal liver cells.* J Exp Med, 1998. **188**(6): p. 1117-24.
9. Giebel, B., *Cell polarity and asymmetric cell division within human hematopoietic stem and progenitor cells.* Cells Tissues Organs, 2008. **188**(1-2): p. 116-26.
10. Chenn, A. and S.K. McConnell, *Cleavage orientation and the asymmetric inheritance of Notch1 immunoreactivity in mammalian neurogenesis.* Cell, 1995. **82**(4): p. 631-41.
11. Gao, F.B. and M. Raff, *Cell size control and a cell-intrinsic maturation program in proliferating oligodendrocyte precursor cells.* J Cell Biol, 1997. **138**(6): p. 1367-77.
12. Hird, S.N., J.E. Paulsen, and S. Strome, *Segregation of germ granules in living Caenorhabditis elegans embryos: cell-type-specific mechanisms for cytoplasmic localisation.* Development, 1996. **122**(4): p. 1303-12.
13. Ikeshima-Kataoka, H., et al., *Miranda directs Prospero to a daughter cell during Drosophila asymmetric divisions.* Nature, 1997. **390**(6660): p. 625-9.
14. Sinclair, D.A. and L. Guarente, *Extrachromosomal rDNA circles--a cause of aging in yeast.* Cell, 1997. **91**(7): p. 1033-42.
15. Giebel, B., *The notch signaling pathway is required to specify muscle progenitor cells in Drosophila.* Mech Dev, 1999. **86**(1-2): p. 137-45.
16. Martinez Arias, A., V. Zecchini, and K. Brennan, *CSL-independent Notch signalling: a checkpoint in cell fate decisions during development?* Curr Opin Genet Dev, 2002. **12**(5): p. 524-33.
17. Osawa, M., et al., *Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell.* Science, 1996. **273**(5272): p. 242-5.
18. Ogawa, M., *Differentiation and proliferation of hematopoietic stem cells.* Blood, 1993. **81**(11): p. 2844-53.

19. Notta, F., et al., *Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment.* Science, 2011. **333**(6039): p. 218-21.
20. Spangrude, G.J., S. Heimfeld, and I.L. Weissman, *Purification and characterization of mouse hematopoietic stem cells.* Science, 1988. **241**(4861): p. 58-62.
21. Weissman, I.L., D.J. Anderson, and F. Gage, *Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations.* Annu Rev Cell Dev Biol, 2001. **17**: p. 387-403.
22. Yang, L., et al., *Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients.* Blood, 2005. **105**(7): p. 2717-23.
23. Oguro, H., L. Ding, and S.J. Morrison, *SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors.* Cell Stem Cell, 2013. **13**(1): p. 102-16.
24. Abramson, S., R.G. Miller, and R.A. Phillips, *The identification in adult bone marrow of pluripotent and restricted stem cells of the myeloid and lymphoid systems.* J Exp Med, 1977. **145**(6): p. 1567-79.
25. Katsura, Y. and H. Kawamoto, *Stepwise lineage restriction of progenitors in lympho-myelopoiesis.* Int Rev Immunol, 2001. **20**(1): p. 1-20.
26. Kawamoto, H., et al., *A map for lineage restriction of progenitors during hematopoiesis: the essence of the myeloid-based model.* Immunol Rev, 2010. **238**(1): p. 23-36.
27. Kondo, M., I.L. Weissman, and K. Akashi, *Identification of clonogenic common lymphoid progenitors in mouse bone marrow.* Cell, 1997. **91**(5): p. 661-72.
28. Akashi, K., et al., *A clonogenic common myeloid progenitor that gives rise to all myeloid lineages.* Nature, 2000. **404**(6774): p. 193-7.
29. Nutt, S.L., et al., *Commitment to the B-lymphoid lineage depends on the transcription factor Pax5.* Nature, 1999. **401**(6753): p. 556-62.
30. Rolink, A.G., C. Schaniel, and F. Melchers, *Stability and plasticity of wild-type and Pax5-deficient precursor B cells.* Immunol Rev, 2002. **187**: p. 87-95.
31. Rolink, A.G., et al., *In vitro and in vivo plasticity of Pax5-deficient pre-B I cells.* Immunol Lett, 2002. **82**(1-2): p. 35-40.
32. Mikkola, I., et al., *Reversion of B cell commitment upon loss of Pax5 expression.* Science, 2002. **297**(5578): p. 110-3.
33. Balciunaite, G., et al., *A B220+ CD117+ CD19- hematopoietic progenitor with potent lymphoid and myeloid developmental potential.* Eur J Immunol, 2005. **35**(7): p. 2019-30.
34. Li, Y.S., et al., *Identification of the earliest B lineage stage in mouse bone marrow.* Immunity, 1996. **5**(6): p. 527-35.
35. Adolfsson, J., et al., *Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment.* Cell, 2005. **121**(2): p. 295-306.
36. Ishikawa, F., et al., *The developmental program of human dendritic cells is operated independently of conventional myeloid and lymphoid pathways.* Blood, 2007. **110**(10): p. 3591-660.

37. Osmond, D.G., A. Rolink, and F. Melchers, *Murine B lymphopoiesis: towards a unified model.* Immunol Today, 1998. **19**(2): p. 65-8.
38. Hardy, R.R., et al., *B-cell commitment, development and selection.* Immunol Rev, 2000. **175**: p. 23-32.
39. Scholz, J.L., M.A. Oropallo, and M.P. Cancro, *Understanding B Cell Biology.* Drugs Targeting B-Cells …, 2014.
40. Ceredig, R., A.G. Rolink, and G. Brown, *Models of haematopoiesis: seeing the wood for the trees.* Nat Rev Immunol, 2009. **9**(4): p. 293-300.
41. Brown, G., et al., *Versatility of stem and progenitor cells and the instructive actions of cytokines on hematopoiesis.* Crit Rev Clin Lab Sci, 2015. **52**(4): p. 168-79.
42. Fogg, D.K., et al., *A clonogenic bone marrow progenitor specific for macrophages and dendritic cells.* Science, 2006. **311**(5757): p. 83-7.
43. Guimaraes, J.E., et al., *Differentiation restriction in the neutrophil-granulocyte, macrophage, eosinophil-granulocyte pathway: analysis by equilibrium density centrifugation.* Leuk Res, 1982. **6**(6): p. 791-800.
44. Altschuler, S.J. and L.F. Wu, *Cellular heterogeneity: do differences make a difference?* Cell, 2010. **141**(4): p. 559-63.
45. Tang, F., et al., *mRNA-Seq whole-transcriptome analysis of a single cell.* Nat Methods, 2009. **6**(5): p. 377-82.
46. Treutlein, B., et al., *Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq.* Nature, 2014. **509**(7500): p. 371-5.
47. Zeisel, A., et al., *Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq.* Science, 2015. **347**(6226): p. 1138-42.
48. Wu, A.R., et al., *Quantitative assessment of single-cell RNA-sequencing methods.* Nat Methods, 2014. **11**(1): p. 41-6.
49. Macosko, E.Z., et al., *Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets.* Cell, 2015. **161**(5): p. 1202-14.
50. Klein, A.M., et al., *Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells.* Cell, 2015. **161**(5): p. 1187-201.
51. Kimmerling, R.J., et al., *A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages.* Nat Commun, 2016. **7**: p. 10220.
52. Chen, A., et al., *On-chip magnetic separation and encapsulation of cells in droplets.* Lab Chip, 2013. **13**(6): p. 1172-81.
53. Fan, H.C., G.K. Fu, and S.P. Fodor, *Expression profiling. Combinatorial labeling of single cells for gene expression cytometry.* Science, 2015. **347**(6222): p. 1258367.
54. Chen, Y., et al., *Continuous enrichment of low-abundance cell samples using standing surface acoustic waves (SSAW).* Lab Chip, 2014. **14**(5): p. 924-30.
55. Bhattacharya, S., et al., *Selective trapping of single mammalian breast cancer cells by insulator-based dielectrophoresis.* Anal Bioanal Chem, 2014. **406**(7): p. 1855-65.
56. Liberale, C., et al., *Integrated microfluidic device for single-cell trapping and spectroscopy.* Sci Rep, 2013. **3**: p. 1258.
57. Min, J.W., et al., *Identification of Distinct Tumor Subpopulations in Lung Adenocarcinoma via Single-Cell RNA-seq.* PLoS One, 2015. **10**(8): p. e0135817.

58. Kim, K.T., et al., *Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells.* Genome Biol, 2015. **16**: p. 127.
59. Chen, X.X. and F. Bai, *Single-cell analyses of circulating tumor cells.* Cancer Biol Med, 2015. **12**(3): p. 184-92.
60. Tang, F., et al., *Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis.* Cell Stem Cell, 2010. **6**(5): p. 468-78.
61. Xue, Z., et al., *Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing.* Nature, 2013. **500**(7464): p. 593-7.
62. Yan, L., et al., *Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells.* Nat Struct Mol Biol, 2013. **20**(9): p. 1131-9.
63. Blainey, P.C. and S.R. Quake, *Dissecting genomic diversity, one cell at a time.* Nat Methods, 2014. **11**(1): p. 19-21.
64. Kolisko, M., et al., *Single-cell transcriptomics for microbial eukaryotes.* Curr Biol, 2014. **24**(22): p. R1081-2.
65. Mahata, B., et al., *Single-cell RNA sequencing reveals T helper cells synthesizing steroids de novo to contribute to immune homeostasis.* Cell Rep, 2014. **7**(4): p. 1130-42.
66. Shalek, A.K., et al., *Single-cell RNA-seq reveals dynamic paracrine control of cellular variation.* Nature, 2014. **510**(7505): p. 363-9.
67. Gren, S.T., et al., *A Single-Cell Gene-Expression Profile Reveals Inter-Cellular Heterogeneity within Human Monocyte Subsets.* PLoS One, 2015. **10**(12): p. e0144351.
68. Jaitin, D.A., et al., *Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types.* Science, 2014. **343**(6172): p. 776-9.
69. Paul, F., et al., *Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.* Cell, 2015. **163**(7): p. 1663-77.
70. Perie, L., et al., *The Branching Point in Erythro-Myeloid Differentiation.* Cell, 2015. **163**(7): p. 1655-62.
71. Stegle, O., S.A. Teichmann, and J.C. Marioni, *Computational and analytical challenges in single-cell transcriptomics.* Nat Rev Genet, 2015. **16**(3): p. 133-45.
72. Metcalf, D., *Hematopoietic cytokines.* Blood, 2008. **111**(2): p. 485-91.
73. Kondo, M., et al., *Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines.* Nature, 2000. **407**(6802): p. 383-6.
74. Rieger, M.A., et al., *Hematopoietic cytokines can instruct lineage choice.* Science, 2009. **325**(5937): p. 217-8.
75. Enver, T., C.M. Heyworth, and T.M. Dexter, *Do stem cells play dice?* Blood, 1998. **92**(2): p. 348-51; discussion 352.
76. Metcalf, D., *Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation.* Blood, 1998. **92**(2): p. 345-7; discussion 352.
77. Endele, M., M. Etzrodt, and T. Schroeder, *Instruction of hematopoietic lineage choice by cytokine signaling.* Exp Cell Res, 2014. **329**(2): p. 207-13.
78. Mackarehtschian, K., et al., *Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors.* Immunity, 1995. **3**(1): p. 147-61.

79. McKenna, H.J., et al., *Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells.* Blood, 2000. **95**(11): p. 3489-97.
80. Peschon, J.J., et al., *Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice.* J Exp Med, 1994. **180**(5): p. 1955-60.
81. von Freeden-Jeffry, U., et al., *Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine.* J Exp Med, 1995. **181**(4): p. 1519-26.
82. Namen, A.E., et al., *Stimulation of B-cell progenitors by cloned murine interleukin-7.* Nature, 1988. **333**(6173): p. 571-3.
83. Conlon, P.J., et al., *Murine thymocytes proliferate in direct response to interleukin-7.* Blood, 1989. **74**(4): p. 1368-73.
84. Goodwin, R.G., et al., *Cloning of the human and murine interleukin-7 receptors: demonstration of a soluble form and homology to a new receptor superfamily.* Cell, 1990. **60**(6): p. 941-51.
85. Maki, K., et al., *Interleukin 7 receptor-deficient mice lack gammadelta T cells.* Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7172-7.
86. Akashi, K., et al., *Bcl-2 rescues T lymphopoiesis in interleukin-7 receptor-deficient mice.* Cell, 1997. **89**(7): p. 1033-41.
87. Maraskovsky, E., et al., *Bcl-2 can rescue T lymphocyte development in interleukin-7 receptor-deficient mice but not in mutant rag-1-/- mice.* Cell, 1997. **89**(7): p. 1011-9.
88. Kondo, M., et al., *Bcl-2 rescues T lymphopoiesis, but not B or NK cell development, in common gamma chain-deficient mice.* Immunity, 1997. **7**(1): p. 155-62.
89. Maraskovsky, E., et al., *Overexpression of Bcl-2 does not rescue impaired B lymphopoiesis in IL-7 receptor-deficient mice but can enhance survival of mature B cells.* Int Immunol, 1998. **10**(9): p. 1367-75.
90. Noguchi, M., et al., *Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans.* Cell, 1993. **73**(1): p. 147-57.
91. Puel, A., et al., *Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency.* Nat Genet, 1998. **20**(4): p. 394-7.
92. Ahsberg, J., et al., *Interleukin-7-induced Stat-5 acts in synergy with Flt-3 signaling to stimulate expansion of hematopoietic progenitor cells.* J Biol Chem, 2010. **285**(47): p. 36275-84.
93. Malin, S., S. McManus, and M. Busslinger, *STAT5 in B cell development and leukemia.* Curr Opin Immunol, 2010. **22**(2): p. 168-76.
94. Malin, S., et al., *Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development.* Nat Immunol, 2010. **11**(2): p. 171-9.
95. Lyman, S.D., et al., *Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase receptor: a proliferative factor for primitive hematopoietic cells.* Cell, 1993. **75**(6): p. 1157-67.
96. Matthews, W., et al., *A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations.* Cell, 1991. **65**(7): p. 1143-52.
97. Gilliland, D.G. and J.D. Griffin, *The roles of FLT3 in hematopoiesis and leukemia.* Blood, 2002. **100**(5): p. 1532-42.

98. Adolfsson, J., et al., *Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity.* Immunity, 2001. **15**(4): p. 659-69.
99. Rasko, J.E., et al., *The flt3/flk-2 ligand: receptor distribution and action on murine haemopoietic cell survival and proliferation.* Leukemia, 1995. **9**(12): p. 2058-66.
100. Holmes, M.L., et al., *Repression of Flt3 by Pax5 is crucial for B-cell lineage commitment.* Genes Dev, 2006. **20**(8): p. 933-8.
101. Ceredig, R., N. Bosco, and A.G. Rolink, *The B lineage potential of thymus settling progenitors is critically dependent on mouse age.* Eur J Immunol, 2007. **37**(3): p. 830-7.
102. Ceredig, R. and A.G. Rolink, *The key role of IL-7 in lymphopoiesis.* Semin Immunol, 2012. **24**(3): p. 159-64.
103. Sambandam, A., et al., *Notch signaling controls the generation and differentiation of early T lineage progenitors.* Nat Immunol, 2005. **6**(7): p. 663-70.
104. Sitnicka, E., et al., *Key role of flt3 ligand in regulation of the common lymphoid progenitor but not in maintenance of the hematopoietic stem cell pool.* Immunity, 2002. **17**(4): p. 463-72.
105. Sitnicka, E., et al., *Complementary signaling through flt3 and interleukin-7 receptor alpha is indispensable for fetal and adult B cell genesis.* J Exp Med, 2003. **198**(10): p. 1495-506.
106. Balciunaite, G., et al., *The role of Notch and IL-7 signaling in early thymocyte proliferation and differentiation.* Eur J Immunol, 2005. **35**(4): p. 1292-300.
107. Veiby, O.P., S.D. Lyman, and S.E. Jacobsen, *Combined signaling through interleukin-7 receptors and flt3 but not c-kit potently and selectively promotes B-cell commitment and differentiation from uncommitted murine bone marrow progenitor cells.* Blood, 1996. **88**(4): p. 1256-65.
108. Ceredig, R., et al., *Increasing Flt3L availability alters composition of a novel bone marrow lymphoid progenitor compartment.* Blood, 2006. **108**(4): p. 1216-22.
109. Tsapogas, P., et al., *In vivo evidence for an instructive role of fms-like tyrosine kinase-3 (FLT3) ligand in hematopoietic development.* Haematologica, 2014. **99**(4): p. 638-46.
110. Shinkai, Y., et al., *RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement.* Cell, 1992. **68**(5): p. 855-67.
111. Vilagos, B., et al., *Essential role of EBF1 in the generation and function of distinct mature B cell types.* J Exp Med, 2012. **209**(4): p. 775-92.
112. Ogawa, M., et al., *B cell ontogeny in murine embryo studied by a culture system with the monolayer of a stromal cell clone, ST2: B cell progenitor develops first in the embryonal body rather than in the yolk sac.* EMBO J, 1988. **7**(5): p. 1337-43.
113. Nakano, T., H. Kodama, and T. Honjo, *Generation of lymphohematopoietic cells from embryonic stem cells in culture.* Science, 1994. **265**(5175): p. 1098-101.
114. Schmitt, T.M. and J.C. Zuniga-Pflucker, *Induction of T cell development from hematopoietic progenitor cells by delta-like-1 in vitro.* Immunity, 2002. **17**(6): p. 749-56.

115. Chomczynski, P. and N. Sacchi, *The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on.* Nat Protoc, 2006. **1**(2): p. 581-5.
116. Chomczynski, P. and N. Sacchi, *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Anal Biochem, 1987. **162**(1): p. 156-9.
117. Borodina, T., J. Adjaye, and M. Sultan, *A strand-specific library preparation protocol for RNA sequencing.* Methods Enzymol, 2011. **500**: p. 79-98.
118. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
119. Au, K.F., et al., *Detection of splice junctions from paired-end RNA-seq data by SpliceMap.* Nucleic Acids Res, 2010. **38**(14): p. 4570-8.
120. Robinson, M.D. and A. Oshlack, *A scaling normalization method for differential expression analysis of RNA-seq data.* Genome Biol, 2010. **11**(3): p. R25.
121. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.
122. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.* Nucleic Acids Res, 2009. **37**(1): p. 1-13.
123. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
124. Reynolds, A.P., et al., *Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms.* Journal of Mathematical Modelling and Algorithms, 2006. **5**(4): p. 475-504.
125. Inlay, M.A., et al., *Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development.* Genes Dev, 2009. **23**(20): p. 2376-81.
126. Mansson, R., et al., *Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity.* Blood, 2010. **115**(13): p. 2601-9.
127. Zhang, J., et al., *Characterization of Siglec-H as a novel endocytic receptor expressed on murine plasmacytoid dendritic cell precursors.* Blood, 2006. **107**(9): p. 3600-8.
128. Blasius, A.L., et al., *Siglec-H is an IPC-specific receptor that modulates type I IFN secretion through DAP12.* Blood, 2006. **107**(6): p. 2474-6.
129. Singh-Jasuja, H., et al., *The mouse dendritic cell marker CD11c is down-regulated upon cell activation through Toll-like receptor triggering.* Immunobiology, 2013. **218**(1): p. 28-39.
130. Nakano, T., *Lymphohematopoietic development from embryonic stem cells in vitro.* Semin Immunol, 1995. **7**(3): p. 197-203.
131. Gao, J., et al., *Characterization of OP9 as authentic mesenchymal stem cell line.* J Genet Genomics, 2010. **37**(7): p. 475-82.
132. Wiktor-Jedrzejczak, W., et al., *Total absence of colony-stimulating factor 1 in the macrophage-deficient osteopetrotic (op/op) mouse.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4828-32.

133. Zuniga-Pflucker, J.C., *T-cell development made simple.* Nat Rev Immunol, 2004. **4**(1): p. 67-72.
134. Koch, U., et al., *Delta-like 4 is the essential, nonredundant ligand for Notch1 during thymic T cell lineage commitment.* J Exp Med, 2008. **205**(11): p. 2515-23.
135. Naik, S.H., et al., *Diverse and heritable lineage imprinting of early haematopoietic progenitors.* Nature, 2013. **496**(7444): p. 229-32.
136. Onai, N., et al., *Identification of clonogenic common Flt3+M-CSFR+ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow.* Nat Immunol, 2007. **8**(11): p. 1207-16.
137. Karsunky, H., et al., *Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells in vivo.* J Exp Med, 2003. **198**(2): p. 305-13.
138. Schlenner, S.M. and H.R. Rodewald, *Early T cell development and the pitfalls of potential.* Trends Immunol, 2010. **31**(8): p. 303-10.
139. Drissen, R., et al., *Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing.* Nat Immunol, 2016. **17**(6): p. 666-76.
140. Lu., Y., et al., *Feature Selection Using Principal Feature Analysis*, in *ACM Multimedia*. 2007.
141. Satija, R. and A.K. Shalek, *Heterogeneity in immune responses: from populations to single cells.* Trends Immunol, 2014. **35**(5): p. 219-29.
142. Dueck, H., J. Eberwine, and J. Kim, *Variation is function: Are single cell differences functionally important?: Testing the hypothesis that single cell variation is required for aggregate function.* Bioessays, 2016. **38**(2): p. 172-80.
143. Bjorklund, A.K., et al., *The heterogeneity of human CD127(+) innate lymphoid cells revealed by single-cell RNA sequencing.* Nat Immunol, 2016. **17**(4): p. 451-60.
144. Guo, G., et al., *Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire.* Cell Stem Cell, 2013. **13**(4): p. 492-505.
145. Kowalczyk, M.S., et al., *Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells.* Genome Res, 2015. **25**(12): p. 1860-72.
146. Miyamoto, T., et al., *Myeloid or lymphoid promiscuity as a critical step in hematopoietic lineage commitment.* Dev Cell, 2002. **3**(1): p. 137-47.
147. Hu, M., et al., *Multilineage gene expression precedes commitment in the hemopoietic system.* Genes Dev, 1997. **11**(6): p. 774-85.
148. Olsson, A., et al., *Single-cell analysis of mixed-lineage states leading to a binary cell fate choice.* Nature, 2016. **537**(7622): p. 698-702.
149. Schlenner, S.M., et al., *Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus.* Immunity, 2010. **32**(3): p. 426-36.
150. Ng, S.Y., et al., *Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells.* Immunity, 2009. **30**(4): p. 493-507.
151. Sakhinia, E., et al., *Gene expression analysis of myeloid and lymphoid lineage markers during mouse haematopoiesis.* Br J Haematol, 2006. **135**(1): p. 105-16.

152. Benz, C., et al., *Hematopoietic stem cell subtypes expand differentially during development and display distinct lymphopoietic programs.* Cell Stem Cell, 2012. **10**(3): p. 273-83.
153. Moignard, V., et al., *Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis.* Nat Cell Biol, 2013. **15**(4): p. 363-72.
154. Ema, H., Y. Morita, and T. Suda, *Heterogeneity and hierarchy of hematopoietic stem cells.* Exp Hematol, 2014. **42**(2): p. 74-82 e2.
155. Moignard, V., et al., *Decoding the regulatory network of early blood development from single-cell gene expression measurements.* Nat Biotechnol, 2015. **33**(3): p. 269-76.
156. Tsang, J.C., et al., *Single-cell transcriptomic reconstruction reveals cell cycle and multi-lineage differentiation defects in Bcl11a-deficient hematopoietic stem cells.* Genome Biol, 2015. **16**: p. 178.
157. Nestorowa, S., et al., *A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation.* Blood, 2016. **128**(8): p. e20-31.
158. Notta, F., et al., *Distinct routes of lineage development reshape the human blood hierarchy across ontogeny.* Science, 2016. **351**(6269): p. aab2116.
159. Tudor, K.S., et al., *Functional assessment of precursors from murine bone marrow suggests a sequence of early B lineage differentiation events.* Immunity, 2000. **12**(3): p. 335-45.
160. Nimmo, R.A., G.E. May, and T. Enver, *Primed and ready: understanding lineage commitment through single cell analysis.* Trends Cell Biol, 2015. **25**(8): p. 459-67.
161. Zandi, S., et al., *Single-cell analysis of early B-lymphocyte development suggests independent regulation of lineage specification and commitment in vivo.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. 15871-6.
162. Dey, S.S., et al., *Integrated genome and transcriptome sequencing of the same cell.* Nat Biotechnol, 2015. **33**(3): p. 285-9.

9. Appendix

Further projects and publications in which I have participated

9.1 Permissive roles of cytokines Interleukin-7 and Flt3-ligand in mouse B cell lineage commitment

Lilly von Muenchow[1], Llucia Albertí-Servera[1], Fabian Klein[1], Giuseppina Capoferri[1], Daniela Finke[2], Rhodri Ceredig[3], Antonius G. Rolink[1] and Panagiotis Tsapogas[1]

[1] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[2] Developmental Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[3] Department of Biosciences, University of Galway, Galway, Ireland

Second author – In press (PNAS)

9.2 Reconstitution of a functional B-cell compartment in immunodeficient mice with pro-B cells propagated with or without stromal cells

Lilly von Muenchow[1], Panagiotis Tsapogas[1], Llucia Albertí-Servera[1], Giuseppina Capoferri[1], Marianne Doelz[1], Hannie Rolink[1], Nabil Bosco[1,2], Rhodri Ceredig[3] and Antonius G. Rolink[1]

[1] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[2] Present address: Nestlé Research Center, Lausanne, Switzerland

[3] Department of Biosciences, University of Galway, Galway, Ireland

Third author – In press (EJI)

9.3 Versatility of stem and progenitor cells and the instructive actions of cytokines shape haematopoiesis

Geoffrey Brown[1], Ciaran J Mooney[1], Llucia Albertí-Servera[2], Lilly von Münchow[2], Kai-Michael Toellner[1], Rhodri Ceredig[3] and Antonius G Rolink[2]

[1] School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[2] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[3] Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland

Third author - Published in Critical Reviews in Clinical Laboratory Sciences, July 2015

9.1 Paper 1:

Permissive roles of cytokines Interleukin-7 and Flt3-ligand in mouse B cell lineage commitment

Lilly von Muenchow[1], Llucia Alberti-Servera[1], Fabian Klein[1], Giuseppina Capoferri[1], Daniela Finke[2], Rhodri Ceredig[3], Antonius Rolink[1] and Panagiotis Tsapogas[1*]

[1] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Mattenstrasse 28, 4054, Basel, Switzerland;

[2] University of Basel Children's Hospital, Basel, Switzerland; Developmental Immunology, Department of Biomedicine, University of Basel, Mattenstrasse 28, 4054, Basel, Switzerland;

[3] National Centre for Biomedical Engineering Science, University of Galway, Galway, Ireland

* Corresponding author contact: Panagiotis Tsapogas panagiotis.tsapogas@unibas.ch

Acknowledgement

A.G.R. is holder of the chair in immunology endowed by L. Hoffmann – La Roche Ltd, Basel. This study was supported by the Swiss National Science Foundation and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement number 315902. R.C. was supported by Science Foundation Ireland under grant numbers SFI09/SRC/B1794 and SFI07/SK/B1233b.

Conflict of interest

The authors declare no commercial or financial conflict of interest.

ABSTRACT

Hematopoietic cells are continuously generated throughout life from hematopoietic stem cells, thus making hematopoiesis a favorable system to study developmental cell lineage commitment. The main factors incorporating environmental signals to developing hematopoietic cells are cytokines, which regulate commitment of hematopoietic progenitors to the different blood lineages by acting either in an instructive or a permissive manner. Flt3-ligand (FL) and Interleukin-7 (IL7) are cytokines pivotal for B cell development, as manifested by the severely compromised B cell development in their absence. However, their precise role in regulating B cell commitment has been the subject of debate. In the present study we assessed the rescue of B cell commitment in mice lacking IL7 but simultaneously over-expressing FL. Results obtained demonstrate that FL over-expression in IL7 deficient mice rescues B cell commitment, resulting in significant Ebf1 and Pax5 expression in $Ly6D^{+}CD135^{+}CD127^{+}CD19^{-}$ precursors and subsequent generation of normal numbers of $CD19^{+}$ B cell progenitors, therefore indicating that IL7 can be dispensable for commitment to the B cell lineage. Further analysis of $Ly6D^{+}CD135^{+}CD127^{+}CD19^{-}$ progenitors in IL7- or FL-deficient mice over-expressing *Bcl2*, as well as in IL7-transgenic mice suggests that both FL and IL7 regulate B cell commitment in a permissive manner; FL by inducing proliferation of $CD135^{+}CD127^{+}Ly6D^{+}CD19^{-}$ progenitors and IL7 by providing survival signals to these progenitors.

INTRODUCTION

Hematopoiesis, the generation of all blood cells from hematopoietic stem cells (HSC), takes place continuously in the adult bone marrow. Accumulating evidence suggests that HSC generate the different hematopoietic lineages via oligo-potent progenitors having limited self-renewal capacity and restricted developmental potentials. Activation of lineage-specific gene transcription in these progenitors eventually leads to their commitment to a particular lineage. Cytokines are the most prominent environmental factors regulating hematopoietic lineage commitment, doing so by acting either in an instructive or a permissive manner (1). In the instructive model, cytokines induce a signaling cascade in progenitors leading to the initiation of a lineage-specific gene program, typically through up-regulation and/or activation of transcription factors, eventually resulting in commitment to a particular lineage. In contrast, the permissive model advocates that commitment of progenitors to different lineages occurs in a cell-autonomous, stochastic manner, with cytokines acting as a selection rather than a commitment factor, promoting the survival and/or proliferation of a specific lineage at the expense of other lineages originating from the same progenitor. Elucidating the precise mode of action of cytokines is technically challenging and therefore the instructive versus permissive roles of cytokines is hotly debated (2-4). Although the permissive model was favored in the past, recent data provide solid evidence for the instructive action of several cytokines including M-CSF, G-CSF, EPO and Flt3-ligand (5-8). However, our understanding of how cytokines regulate hematopoiesis remains elusive, as different cytokines can act in various ways and their function might be cell-context dependent (9). Moreover, most studies to date have addressed cytokine regulated myeloid differentiation with relatively little information on lymphoid lineage commitment.

That Interleukin-7 (IL7) is a crucial cytokine for B cell generation is demonstrated by the dramatic defect in B cell generation in mice lacking either the cytokine (10) or its receptor (11). Interestingly, while human B cell progenitors are also responsive to IL7 (12), disruption of IL7 signaling caused by mutations does not ablate B cell development in man (13, 14). IL7 was initially identified as a growth factor for B cell progenitors (15) and early studies demonstrated that *in vivo* over-

expression of the pro-survival gene *Bcl2* did not rescue B cell development in the absence of IL7 signaling suggesting that IL7 acts in an instructive manner in B cell commitment (16, 17). The subsequent findings that uncommitted Common Lymphoid Progenitors (CLP) from *Il7*$^{-/-}$ mice lacked expression of the transcription factor Early B-cell Factor 1 (Ebf1) (18) and that Ebf1 over-expression partially restored B cell generation from these CLP (19), led to the hypothesis that IL7, through Stat5 activation, instructs commitment to the B cell lineage by initiating Ebf1 expression in uncommitted progenitors. Supporting this hypothesis, a putative Stat5 binding site was later identified in one of the *Ebf1* promoters (20). However, a more recent study has shown that Bcl2 can rescue B cell generation in a *Stat5* conditional knock-out mouse (21). Furthermore, the Ebf1-expressing fraction of CLP (Ly6D^{+} CLP) is dramatically reduced in *Il7*$^{-/-}$ mice (22), therefore providing an alternative possibility for the reduced Ebf1 expression observed in *Il7*$^{-/-}$ CLP. Interestingly, B cell lineage commitment is initiated at the molecular level in Ly6D^{+}CD19^{-} progenitors (23). Hence, while the importance of IL7 as a growth factor for committed B cell progenitors has been well established, it remains unclear whether it instructs oligo-potent progenitors to commit to the B cell lineage through *Ebf1* and *Pax5* up-regulation.

Ftl3-ligand (FL), the only known ligand for the Flt3 receptor (CD135), is a cytokine important for the generation of many hematopoietic lineages and its function has gained much attention as mutations in FL signaling are commonly found in Acute Myeloid Leukemias (AML) (24). Committed B cell progenitors do not express CD135 since expression of the B cell commitment factor Pax5 leads to its down-regulation (25). However, upon transplantation, bone marrow progenitors from *Flt3*$^{-/-}$ and *Flt3l*$^{-/-}$ mice reconstitute the B cell compartment poorly (26, 27), and FL was found to be essential for maintaining normal numbers of uncommitted B cell progenitors (28).

Recently, we described a FL-transgenic mouse model (hereafter *Flt3Ltg*) expressing high levels of FL *in vivo*, which has enabled us to suggest an instructive role for FL in early stages of hematopoiesis (8). By breeding these mice with *Il7*$^{-/-}$ mice we herein show that increased FL levels can rescue B cell commitment in CD135^{+}CD127^{+}CD19^{-} progenitors and restore early CD19^{+} B cell progenitor numbers

in the absence of IL7 signaling, suggesting a permissive role for IL7 in B cell commitment. Further analyses of a combination of mouse genotypes over-expressing or lacking FL and IL7, as well as the pro-survival gene *Bcl-2*, have enabled us to identify a permissive role for both IL7 and FL in B cell commitment.

RESULTS

Increased *in vivo* levels of FL rescue B cell commitment in *Il7*$^{-/-}$ Ly6D^{+}CD19^{-} progenitors

We have previously characterized an uncommitted B cell progenitor with combined lymphoid and myeloid potential (Early Progenitor with Lymphoid and Myeloid potential – EPLM) (29). EPLM can be further subdivided by SiglecH, CD11c, CD115 and Ly6D expression enabling us to identify the Ly6D^{+}SiglecH^{-}CD11c^{-}CD115^{-} fraction of EPLM (hereafter Ly6D^{+} EPLM) as the population containing most B cell potential, while being devoid of myeloid potential (Fig. S1B). This EPLM subpopulation is identified as Lin^{-}CD19^{-}CD117^{int}B220^{int}Ly6D^{+}CD135^{+}CD127^{+} (Figure 1A) therefore partially overlapping phenotypically with Ly6D^{+} CLP (Fig. S1A) and pre-pro-B cells (30, 31). Ly6D^{+} EPLM numbers in *Il7*$^{-/-}$ and *Flt3l*$^{-/-}$ mice are significantly decreased compared to WT; 7-fold for *Il7*$^{-/-}$ and 13-fold for *Flt3l*$^{-/-}$ respectively and a similar dramatic decrease was observed in Ly6D^{+} CLP from both mutant mice (Figure 1B-C). FL deficiency also affected the numbers of Ly6D^{-} EPLM and CLP, while IL7 did not (Fig. S1C-D). Therefore, Ly6D^{+} EPLM/CLP represent the earliest developmental stage of the B cell pathway affected by the absence of IL7.

We have recently generated a mouse model expressing high *in vivo* levels of FL (8). The progenitor compartment of these mice showed a dramatic increase in EPLM and CLP numbers, with their Ly6D^{+} fractions increased 90-fold and 28-fold respectively relative to WT (Figure 1D-E). We crossed *Flt3Ltg* with *Il7*$^{-/-}$ mice to assess the extent to which increased FL levels could potentially rescue the loss of Ly6D^{+}CD19^{-} progenitors in *Il7*$^{-/-}$ mice. As shown in Figure 2 A-B, *in vivo* over-expression of FL leads to a significant increase in *Flt3Ltg-Il7*$^{-/-}$ EPLM and CLP numbers, reaching levels of those in *Flt3Ltg* mice. Crucially, a full rescue of Ly6D^{+} EPLM and CLP can be seen in these mice, with a striking 470-fold and 31-fold increase in numbers compared to their *Il7*$^{-/-}$ counterparts (Figure 2A-B). Furthermore, the numbers of the earliest committed CD19^{+}CD117^{+} pro-B cells were fully restored in *Flt3Ltg-Il7*$^{-/-}$ mice, showing a 251-fold increase compared to *Il7*$^{-/-}$ (Figure 2C and Fig. S2). However, this rescue was less pronounced in downstream CD19^{+}CD117^{-}IgM^{-} and CD19^{+}IgM^{+} B cell stages, since these cells require IL7 to

expand. As a consequence of this rescue in bone marrow B cell development, numbers of splenic marginal zone and follicular B cells were significantly increased in *Flt3Ltg-Il7*$^{-/-}$ mice compared to *Il7*$^{-/-}$ (Figure 2D). While thymic T cell development was not rescued in *Flt3Ltg-Il7*$^{-/-}$ mice (Fig S3), a significant increase in splenic T cell numbers was observed (Fig. S4) as a result of their expansion upon FL over-expression (32).

To assess whether these rescued *Flt3Ltg-Il7*$^{-/-}$ Ly6D^{+}CD19^{-} progenitors could give rise to B cells *in vitro,* we sorted *Flt3Ltg-Il7*$^{-/-}$ Ly6D^{+} EPLM and plated them at limiting dilution on OP9 stromal cells in the presence of IL7. As shown in Figure 3A, *Flt3Ltg-Il7*$^{-/-}$ Ly6D^{+} EPLM could generate B cells at similar frequencies to their WT and *Flt3Ltg* counterparts, whereas the few *Il7*$^{-/-}$ Ly6D^{+} EPLM isolated could not. A rescue in Ly6D^{+} EPLM was also observed when *Il7*$^{-/-}$ mice were injected with FL (Figure 3E) and when plated under the same conditions these rescued Ly6D^{+} EPLM also showed a restored *in vitro* B cell potential (Fig. S5A). Further, when transplanted into irradiated *Rag2*$^{-/-}$ mice they were able to generate IgM^{+} B cells (Fig S5B-C). Thus, increased FL levels restore the generation of Ly6D^{+} progenitors, rather than merely expanding the few Ly6D^{+} EPLM/CLP found in *Il7*$^{-/-}$ mice. Real-time quantitative PCR analysis of Ly6D^{+} EPLM from *Flt3Ltg-Il7*$^{-/-}$ mice revealed significant expression of *Ebf1*, *Pax5* and *Foxo1* transcription factors' mRNA in the absence of IL7 (Figure 3B). Ebf1 expression at the protein level was confirmed by intracellular FACS staining (Figure 3C-D). Even though the percentage of *Flt3Ltg-Il7*$^{-/-}$ Ebf1^{+}Ly6D^{+} EPLM did not reach WT levels, it was similar to the one found in *Flt3Ltg* mice, which produce IL7. Therefore, *Ebf1/Pax5* expression and subsequent commitment to the B cell fate can occur in the absence of IL7 signaling, arguing against an instructive role of this cytokine in B cell commitment.

CD127 (IL7Rα) is a receptor shared between IL7 and thymic stromal lymphopoietin (TSLP), a cytokine capable of rescuing B cell development when over-expressed in the absence of IL7 (33). Since TSLP is produced by dendritic cells (34), which are dramatically expanded in *Flt3Ltg* mice (8), *in vivo* FL over-expression could lead to increased levels of TSLP thereby rescuing B cell development in *Flt3Ltg-Il7*$^{-/-}$ mice. To investigate this possibility we injected *Il7*$^{-/-}$ or *Il7rα*$^{-/-}$ mice with FL as described above and assessed the rescue of Ly6D^{+} EPLM and downstream CD19^{+}

progenitors. FL injections into *Il7*$^{-/-}$ mice resulted in a significant increase in Ly6D^{+} EPLM and CD19^{+}CD117^{+} B cell progenitors, comparable to the rescue observed in *Flt3Ltg-Il7*$^{-/-}$ mice (Figure 3E). FL injected *Il7rα*$^{-/-}$ mice also demonstrated a significant rescue of Ly6D^{+} EPLM and CD19^{+}CD117^{+} pro-B cells, indicating that the observed rescue of B cell commitment in *Flt3Ltg-Il7*$^{-/-}$ mice is not mediated through the action of TSLP.

IL7 promotes survival, but not proliferation, of Ly6D^{+}CD135^{+}CD127^{+}CD19^{-} progenitors

Even though our *Flt3Ltg-Il7*$^{-/-}$ mouse model suggests that IL7 is dispensable for B cell commitment, the dramatic decrease in *Il7*$^{-/-}$ Ly6D^{+} EPLM/CLP argues for a role of IL7 in the maintenance of this population when FL levels are limiting, by promoting either their survival or their proliferation. To investigate the potential role of IL7 as a survival factor for Ly6D^{+}CD135^{+}CD127^{+}CD19^{-} progenitors, we crossed *Il7*$^{-/-}$ mice with mice expressing the pro-survival gene *Bcl2* (35). *Bcl2*tg-*Il7*$^{-/-}$ mice demonstrated a minor but statistically significant 2.6-fold increase in Ly6D^{+} EPLM and 2.2-fold increase in Ly6D^{+} CLP numbers compared to *Il7*$^{-/-}$ mice (Figure 4A-B). Cell cycle stage analysis of Ly6D^{+} EPLM of these mice indicated that most of the cells rescued by *Bcl2* are in a quiescent state (Fig. S6) and do not proliferate in response to cytokines, thereby compromising to some extent the rescue of these progenitors' numbers. Importantly, when plated on OP9 stromal cells plus IL7, *Bcl2*tg-*Il7*$^{-/-}$ Ly6D^{+} EPLM generated B cells at frequencies similar to WT mice (Figure 4C), indicating that these rescued Ly6D^{+} cells had B cell potential. Indeed, when analyzing bone marrow CD19^{+} committed progenitors, we could see a significant 68-fold increase in the earliest CD19^{+}CD117^{+} pro-B cell compartment, compared to *Il7*$^{-/-}$ (Figure 4D). Due to the their quiescent state {Fig. S6 and (36)} and the IL7 dependence of their proliferation, *Bcl2*tg-*Il7*$^{-/-}$ CD19^{+}CD117^{+} numbers did not reach WT levels, whereas downstream CD19^{+} immature B cells showed a less pronounced, but significant rescue (Figure 4D). In the periphery of these mice, marginal zone and follicular B cell numbers were increased, whereas as previously reported (16), T cell numbers were rescued (Fig. S7). Therefore, providing an extra *Bcl2*-mediated survival signal *in vivo* partially rescues *Il7*$^{-/-}$ Ly6D^{+}CD19^{-} progenitors with B cell potential and restores

significantly the generation of $CD19^+$ progenitors. This suggests a role for IL7 in facilitating the survival of $Ly6D^+CD135^+CD127^+CD19^-$ progenitors.

To evaluate the potential proliferative effect of IL7 on $Ly6D^+CD135^+CD127^+CD19^-$ progenitors, we analyzed a transgenic mouse model in which *Il7* expression is driven by an MHC Class II promoter, resulting in increased *in vivo* levels of IL7 (37). These mice exhibit a lymphoproliferative phenotype with increased numbers of $CD19^+$ B cells (38). In contrast to bone marrow $CD19^+$ cells, $Ly6D^+$ EPLM numbers did not increase in response to elevated IL7 (Figure 5A-C). In addition, the cell cycle profile of $Ly6D^+$ EPLM remained unaltered in *Il7*tg mice compared to WT (Figure 5D), arguing against a proliferative action of IL7 on these progenitors. To exclude the possibility that a proliferative signal by FL present in these mice compromised the effect of increased IL7 on the cell cycle status of $Ly6D^+$ EPLM, we crossed *Il7*tg with *Flt3l*$^{-/-}$ mice. Over-expression of IL7 *in vivo* did not result in a significant increase in $Ly6D^+$ EPLM or CLP numbers in the absence of FL (Figure 5E and FigS8). In contrast, a 3-fold increase in $CD19^+CD117^+$ numbers was observed (Figure 5F), in agreement with the proliferative effect of IL7 on $CD19^+$ B cells. This resulted in a small, but significant, increase in splenic follicular B cells (Fig. S9). Moreover, cell cycle analysis of *Il7*tg-*Flt3l*$^{-/-}$ $Ly6D^+$ EPLM showed no significant change in their cycling profile compared to their *Flt3l*$^{-/-}$ counterparts (Figure 5G). Therefore, we conclude that while IL7 acts as a proliferative factor for $CD19^+$ committed B cells, it does not do so for their $Ly6D^+CD135^+CD127^+CD19^-$ precursors.

FL induces proliferation of $Ly6D^+CD135^+CD127^+CD19^-$ progenitors

As evident in Figure 5G, loss of *in vivo* FL signaling affected the proliferative status of $Ly6D^+$ EPLM. Comparison of $Ly6D^+$ EPLM numbers in mice either lacking or over-expressing FL showed a 14-fold reduction in *Flt3l*$^{-/-}$ $Ly6D^+$ EPLM numbers compared to WT, while *Flt3Ltg* $Ly6D^+$ EPLM increased 105-fold (Figure 6A). A similar response to FL levels was observed for $Ly6D^+$ CLP (Figure 6A). Cell cycle analysis of $Ly6D^+$ EPLM from these mice showed a significant increase in the percentage of $Ki67^-$ $DAPI^-$ cells and a decrease in the percentage of $Ki67^+$ cells when FL signaling was absent, while *Flt3Ltg* $Ly6D^+$ EPLM showed the reverse (Figure 6B and Fig. S10). Thus,

our data indicate that FL promotes the proliferation of Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ progenitors.

To evaluate if FL also regulates the survival of Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ progenitors, we crossed *Flt3l*$^{-/-}$ mice with *Bcl2*tg mice. Thus, *Bcl2*tg-*Flt3l*$^{-/-}$ mice showed a minor 2-fold increase in Ly6D$^+$ EPLM numbers compared to their *Flt3l*$^{-/-}$ counterparts (1.8-fold for Ly6D$^+$ CLP) (Figure 6C). Nevertheless, the *in vitro* B cell potential of *Flt3l*$^{-/-}$ Ly6D$^+$ EPLM progenitors was not improved by *Bcl2* over-expression (Figure 6D). Downstream CD19$^+$ progenitors also demonstrated a partial, but significant, rescue (Figure 6E). Our analysis of *Bcl2*tg-*Flt3l*$^{-/-}$ mice suggests that the reduction in Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ progenitors observed in *Flt3l*$^{-/-}$ mice can only be partially explained by a survival role of FL. In contrast, the clear change in the numbers and cycling profile of these progenitors in response to the absence or over-abundance of FL *in vivo*, as well as the inability of Bcl2 to rescue their *in vitro* B cell potential, points towards proliferation as being the main effector function of FL at this developmental stage.

FL does not instruct commitment to the B cell lineage

The striking rescue in B cell commitment observed in our *Flt3Ltg-Il7*$^{-/-}$ mice could be explained by a potential instructive role of FL when present at high levels *in vivo*. However, increased FL did not result in *Ebf1* or *Pax5* up-regulation (Figure 3B-D). Moreover, analysis of *Flt3l*$^{-/-}$ Ly6D$^+$ EPLM showed that while absence of FL *in vivo* leads to a reduction in the numbers of Ly6D$^+$ EPLM (Figure 1C), it does not significantly reduce the percentage of Ebf1$^+$ cells within the population (Figure 7A-B), consistent with a permissive rather than instructive role of FL. Finally, the decrease in the Ebf1$^+$ fraction of Ly6D$^+$ EPLM upon exposure to high levels of FL was reflected in the increased ability of these progenitors to give rise to T cells *in vitro*, as manifested by the high frequency of T cell clone generation when *Flt3Ltg* Ly6D$^+$ EPLM were plated on OP9DL1 stromal cells in the presence of IL7 (Figure 7C). The above data suggest that FL does not instruct commitment to the B cell lineage through up-regulation of *Ebf1* and *Pax5* expression.

DISCUSSION

Commitment to the B cell lineage is mediated by the expression of Ebf1 and Pax5 transcription factors and it is initiated in $CD135^{+}CD127^{+}Ly6D^{+}$ progenitors prior to CD19 expression (30, 31). In *Il7*$^{-/-}$ mice, this $Ly6D^{+}$ CLP compartment is significantly reduced (22), a finding confirmed in the present study for both CLP and EPLM, a $B220^{int/+}$ population partly overlapping with CLP and pre-pro B cells (Figure 1B-C). The proliferative effect of IL7 on committed $CD19^{+}$ B cell progenitors (38) makes the investigation of its role in B cell commitment challenging when using $CD19^{+}$ cells as readout. Hence, we assessed the role of IL7 in B cell commitment by analyzing the $Ly6D^{+}$ CLP/EPLM compartment in different mouse models. Our analysis of *Flt3Ltg-Il7*$^{-/-}$mice showed a complete rescue of $Ly6D^{+}$ CLP/EPLM numbers *in vivo* and their B cell potential *in vitro* and *in vivo,* while Ebf1 and Pax5 were expressed at similar levels to *Flt3Ltg* mice, thereby indicating that IL7 signaling is not required for their up-regulation at the $Ly6D^{+}CD19^{-}$ stage (Figures 2 and 3). These results suggest that IL7 is not acting as an instructive cytokine in B cell commitment by initiating Ebf1 and Pax5 expression at the $CD135^{+}CD127^{+}CD19^{-}$ stage, as previously hypothesized (18-20), but rather as a permissive one.

Early investigations had shown that *Bcl2* over-expression in the absence of IL7 signaling could rescue T cell (39, 40) but not B cell development (16, 17). However, a more recent study demonstrated a *Bcl2*-mediated rescue of $CD19^{+}$ progenitors in conditional *Stat5*$^{-/-}$ mice, as well as a strong activation of the pro-survival gene *Mcl1* expression by Stat5 (21), therefore suggesting a survival role for IL7 in B cell development. Our use of *Il7*$^{-/-}$ mice instead of *Il7rα*$^{-/-}$, which allows the assessment of progenitor *in vitro* B cell potential, and our focus on $Ly6D^{+}CD135^{+}CD127^{+}CD19^{-}$ progenitors, has enabled us to confirm the latter findings and extend them to the $CD19^{-}$ stage where B cell commitment events are initiated at the molecular level. Interestingly, *Il7*tg mice analysis showed that IL7 indeed acts as a proliferative factor for committed $CD19^{+}$ cells, but not for their $CD19^{-}$ precursors. Even in the absence of FL, excess IL7 was unable to significantly increase $Ly6D^{+}$ CLP/EPLM numbers, while it did so for $CD19^{+}$ B cell progenitors (Figure 5). Hence, we propose that the main role of IL7 at the $CD135^{+}CD127^{+}CD19^{-}$ stage is to provide

survival signals to the progenitors until they commit to the B cell lineage upon Pax5 and CD19 expression, after which it additionally induces their proliferation (Figure 7D). This survival role becomes particularly critical when FL levels are limiting, thereby explaining the reduction in $Ly6D^+$ CLP/EPLM seen in *Il7*$^{-/-}$ mice. Our study, in agreement with previous data (21), identifies a common, permissive rather than instructive role for IL7 in both B and T cell development (39, 40).

The rescue in B cell commitment without active IL7 signaling occurs when FL is expressed above physiological levels. Even though a minor role for FL as a survival factor for $CD135^+CD127^+Ly6D^+CD19^-$ progenitors cannot be excluded, the main effect of FL on these progenitors seems to be the induction of their proliferation, as suggested by their expansion and their increased cycling upon FL over-expression, with the reverse phenotype observed upon loss of FL signaling (Figure 6). Moreover, increased FL leads to expansion of $Lin^-CD117^+Sca1^+$ cells (LSK) (8), thereby increasing the developmental input into the $CD135^+CD127^+Ly6D^+CD19^-$ progenitor stage. None of the mouse models analyzed in the present study gave any evidence for an instructive role of FL in B cell commitment. In contrast, excess FL resulted in a proportional reduction of Ebf- and Pax5-expressing $Ly6D^+CD19^-$ progenitors (Figures 3 and 7). One explanation for this reduction could be the increased percentage of cycling *Flt3Ltg* $Ly6D^+CD19^-$ progenitors, resulting in a decreased fraction initiating the B cell developmental program. Alternatively, another environmental factor, responsible for initiation of *Ebf1/Pax5* expression and B cell commitment, could be the limiting factor in *Flt3Ltg* mice, thus leading to a smaller fraction of the expanded $Ly6D^+CD19^-$ compartment entering the B cell pathway. Our conclusion is that FL is mainly responsible for generating enough $CD135^+CD127^+Ly6D^+CD19^-$ progenitors, both by inducing their proliferation and by increasing their developmental input from the LSK compartment (Figure 7D) (41, 42). As a result, increased levels of FL in *Flt3Ltg-Il7*$^{-/-}$ mice lead to a dramatic increase in $CD135^+CD127^+Ly6D^+CD19^-$ progenitor numbers, therefore surpassing the need for the survival role of IL7 at this stage and resulting in a sufficient fraction of them committing to the B cell lineage.

The generation of B cell progenitors in *Flt3Ltg-Il7*$^{-/-}$ mice is reminiscent of the apparent IL7 independency of human B lymphopoiesis, where relatively normal numbers of B cells are seen in patients with mutations in components of the IL7

signaling pathway (13, 14). However, all patients with such mutations are neonates and in neonatal *Il7*$^{-/-}$ mice, B cell development also takes place (43). Therefore, the apparent difference in the IL7 dependency of B cell development between man and mouse could actually reflect the corresponding difference between fetal/neonatal and adult lymphopoiesis. Our data showing that increased FL signaling can rescue B cell commitment in the absence of IL7 could provide a potential explanation for this difference. Fetal/neonatal CD135^{+}CD127^{+}CD19^{-} progenitors might be exposed to higher levels of FL and/or show higher sensitivity to FL signaling than adult CD135^{+}CD127^{+}CD19^{-} progenitors. Indeed, previous studies showed that despite a preferable response of fetal B cell progenitors to TSLP, FL signaling remains an absolute requirement for fetal B lymphopoiesis (44, 45).

The instructive or permissive progenitor regulation of lineage commitment by cytokines is a complex process, in which cytokines can initiate developmental transcription programs in progenitors. However, the reverse is also true, since the particular epigenetic, transcriptional and signaling landscape of a cell can affect its response to a cytokine (9). Indeed, while previous analysis of *Flt3Ltg* mice indicated an instructive role for FL in promoting differentiation of multi-potent progenitors towards lympho-myeloid and away from erythroid fate (8), our present data show that FL acts in a permissive manner for B cell commitment of CD135^{+}CD127^{+}CD19^{-} progenitors. In addition, whereas IL7 induces proliferation of committed CD19^{+} B cell progenitors, it does not do so on CD127^{+}CD19^{-} progenitors, suggesting that upon commitment to the B cell lineage, changes in the transcription factor and intracellular signaling landscape influence the effector function of IL7. Therefore, our present data further support the notion of a cell-context dependent cytokine action.

The *Ebf1/Pax5* up-regulation and subsequent B cell commitment in *Flt3Ltg-Il7*$^{-/-}$ mice shown herein raises the issue of the potential extra-cellular regulation of B cell commitment. One possibility could be that another environmental signal from the bone marrow microenvironment - other than IL7, TSLP and FL - initiates *Ebf1* expression in CD135^{+}CD127^{+}Ly6D^{+}CD19^{-} progenitors resulting in Pax5/CD19 expression and B cell commitment. Alternatively, as yet uncommitted CD135^{+}CD127^{+}Ly6D^{+}CD19^{-} progenitors could express Ebf1 in a cell-autonomous, stochastic, manner with some obtaining sufficient Ebf1 to initiate the B cell gene

program and eventually commit to the B cell lineage. The intricate transcription factor network sustaining B cell commitment through a series of positive feedback regulatory loops (46) provides conceptual support for the latter hypothesis.

MATERIALS AND METHODS

Mice

For breeding and analysis, age- and sex-matched C57BL/6 *Flt3l*$^{-/-}$ (27), *Flt3Ltg* (8), *Il7*$^{-/-}$ (10), *Il7rα*$^{-/-}$ (11), *Il7*tg (38), and (C57BL/6 x C3H) *Bcl2*tg (35) mice backcrossed with C57BL/6 for at least 5 generations were used at 6–11 weeks of age. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. Animal experiments were carried out within institutional guidelines (authorization number 1888 from cantonal veterinarian office, Basel).

Antibodies, flow cytometry, and sorting

For analysis, cells were flushed from femurs of the two hind legs of mice. The procedure was performed in PBS containing 0.5% BSA and 5mM EDTA. For detection of Ebf1 and cell cycle analysis, cells were fixed and permeabilized after cell-surface staining using the Foxp3 Fix/Perm buffer set (eBioscience)**,** and subsequently stained with PE-conjugated anti-Ebf1 (T26-818) or FITC-conjugated anti-Ki67 (B56) and DAPI, according to the supplier's protocol. Flow cytometry was done using a BD LSRFortessa (BD Biosciences) and data were analyzed using FlowJo Software (Treestar). For cell sorting, a FACSAria IIu (BD Biosciences) was used (>98% purity).

In vitro limiting dilution assays

Experiments have been performed as previously described (47). Briefly, OP9 or OP9DL1 stromal cells were plated on flat-bottom 96-well plates one day before the initiation of co-cultures, at a concentration of 3000 cells per well. The following day stromal cells were γ-irradiated (3000 rad) and the sorted progenitor cells were added at different concentrations. Cultures were maintained in IMDM medium supplemented with 5×10^{-5} M β-mercaptoethanol, 1 mM glutamine, 0.03% (wt/vol) primatone, 100 U/mL penicillin, 100 μg/mL streptomycin, 5% FBS and 10% IL7-conditioned medium. After 14 days in culture all wells were inspected under an inverted microscope and wells containing colonies of more than 50 cells were scored as positive.

Quantitative real-time PCR

RNA extraction was performed using TRI Reagent® (Life Technologies) followed by cDNA synthesis using using GoScript™ Reverse Transcriptase (Promega). Real-time PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems).

Statistical analysis

Statistical analysis was performed with Prism 6.0g software (GraphPad Software, Inc.). Two-tailed unpaired Student *t* tests were used for statistical comparisons. If not differently indicated, data are presented as mean values ± SD or SEM. n.s. not significant or $P > 0.05$, *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$.

AUTHOR CONTRIBUTIONS

L.vM. performed experiments, analyzed data and revised the manuscript; L.A-S., F.K. and G.C. performed experiments; D.F. provided *Il7*$^{-/-}$ and *Il7rα*$^{-/-}$ mice and revised the manuscript; R.C provided *Il7*tg mice and revised the manuscript; A.R. designed experiments, analyzed data, revised the manuscript and supervised the project; P.T. designed and performed experiments, analyzed data, wrote the manuscript and supervised the project.

REFERENCES

1. Metcalf D (2008) Hematopoietic cytokines. *Blood* 111(2):485-491.
2. Endele M, Etzrodt M, & Schroeder T (2014) Instruction of hematopoietic lineage choice by cytokine signaling. *Experimental cell research* 329(2):207-213.
3. Enver T, Heyworth CM, & Dexter TM (1998) Do stem cells play dice? *Blood* 92(2):348-351; discussion 352.
4. Metcalf D (1998) Lineage commitment and maturation in hematopoietic cells: the case for extrinsic regulation. *Blood* 92(2):345-347; discussion 352.
5. Grover A*, et al.* (2014) Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *The Journal of experimental medicine* 211(2):181-188.
6. Mossadegh-Keller N*, et al.* (2013) M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* 497(7448):239-243.
7. Rieger MA, Hoppe PS, Smejkal BM, Eitelhuber AC, & Schroeder T (2009) Hematopoietic cytokines can instruct lineage choice. *Science* 325(5937):217-218.
8. Tsapogas P*, et al.* (2014) In vivo evidence for an instructive role of fms-like tyrosine kinase-3 (FLT3) ligand in hematopoietic development. *Haematologica* 99(4):638-646.
9. Sarrazin S & Sieweke M (2011) Integration of cytokine and transcription factor signals in hematopoietic stem cell commitment. *Seminars in immunology* 23(5):326-334.
10. von Freeden-Jeffry U*, et al.* (1995) Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *The Journal of experimental medicine* 181(4):1519-1526.
11. Peschon JJ*, et al.* (1994) Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *The Journal of experimental medicine* 180(5):1955-1960.
12. Parrish YK*, et al.* (2009) IL-7 Dependence in human B lymphopoiesis increases during progression of ontogeny from cord blood to bone marrow. *Journal of immunology* 182(7):4255-4266.
13. Noguchi M*, et al.* (1993) Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans. *Cell* 73(1):147-157.
14. Puel A, Ziegler SF, Buckley RH, & Leonard WJ (1998) Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nature genetics* 20(4):394-397.
15. Namen AE*, et al.* (1988) Stimulation of B-cell progenitors by cloned murine interleukin-7. *Nature* 333(6173):571-573.
16. Kondo M, Akashi K, Domen J, Sugamura K, & Weissman IL (1997) Bcl-2 rescues T lymphopoiesis, but not B or NK cell development, in common gamma chain-deficient mice. *Immunity* 7(1):155-162.
17. Maraskovsky E, Peschon JJ, McKenna H, Teepe M, & Strasser A (1998) Overexpression of Bcl-2 does not rescue impaired B lymphopoiesis in IL-7

receptor-deficient mice but can enhance survival of mature B cells. *International immunology* 10(9):1367-1375.
18. Dias S, Silva H, Jr., Cumano A, & Vieira P (2005) Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. *The Journal of experimental medicine* 201(6):971-979.
19. Kikuchi K, Lai AY, Hsu CL, & Kondo M (2005) IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *The Journal of experimental medicine* 201(8):1197-1203.
20. Roessler S*, et al.* (2007) Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Molecular and cellular biology* 27(2):579-594.
21. Malin S*, et al.* (2010) Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. *Nature immunology* 11(2):171-179.
22. Tsapogas P*, et al.* (2011) IL-7 mediates Ebf-1-dependent lineage restriction in early lymphoid progenitors. *Blood* 118(5):1283-1290.
23. Mansson R*, et al.* (2008) B-lineage commitment prior to surface expression of B220 and CD19 on hematopoietic progenitor cells. *Blood* 112(4):1048-1055.
24. Gilliland DG & Griffin JD (2002) The roles of FLT3 in hematopoiesis and leukemia. *Blood* 100(5):1532-1542.
25. Holmes ML, Carotta S, Corcoran LM, & Nutt SL (2006) Repression of Flt3 by Pax5 is crucial for B-cell lineage commitment. *Genes & development* 20(8):933-938.
26. Mackarehtschian K*, et al.* (1995) Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity* 3(1):147-161.
27. McKenna HJ*, et al.* (2000) Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* 95(11):3489-3497.
28. Sitnicka E*, et al.* (2002) Key role of flt3 ligand in regulation of the common lymphoid progenitor but not in maintenance of the hematopoietic stem cell pool. *Immunity* 17(4):463-472.
29. Balciunaite G, Ceredig R, Massa S, & Rolink AG (2005) A B220+ CD117+ CD19- hematopoietic progenitor with potent lymphoid and myeloid developmental potential. *European journal of immunology* 35(7):2019-2030.
30. Inlay MA*, et al.* (2009) Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes & development* 23(20):2376-2381.
31. Mansson R*, et al.* (2010) Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity. *Blood* 115(13):2601-2609.
32. Swee LK, Bosco N, Malissen B, Ceredig R, & Rolink A (2009) Expansion of peripheral naturally occurring T regulatory cells by Fms-like tyrosine kinase 3 ligand treatment. *Blood* 113(25):6277-6287.
33. Chappaz S, Flueck L, Farr AG, Rolink AG, & Finke D (2007) Increased TSLP availability restores T- and B-cell compartments in adult IL-7 deficient mice. *Blood* 110(12):3862-3870.

34. Kashyap M, Rochman Y, Spolski R, Samsel L, & Leonard WJ (2011) Thymic stromal lymphopoietin is produced by dendritic cells. *Journal of immunology* 187(3):1207-1211.
35. Domen J, Gandy KL, & Weissman IL (1998) Systemic overexpression of BCL-2 in the hematopoietic system protects transgenic mice from the consequences of lethal irradiation. *Blood* 91(7):2272-2282.
36. O'Reilly LA, Huang DC, & Strasser A (1996) The cell death inhibitor Bcl-2 and its homologues influence control of cell cycle entry. *The EMBO journal* 15(24):6979-6990.
37. Fisher AG*, et al.* (1993) Lymphoproliferative disorders in an IL-7 transgenic mouse line. *Leukemia* 7 Suppl 2:S66-68.
38. Mertsching E, Grawunder U, Meyer V, Rolink T, & Ceredig R (1996) Phenotypic and functional analysis of B lymphopoiesis in interleukin-7-transgenic mice: expansion of pro/pre-B cell number and persistence of B lymphocyte development in lymph nodes and spleen. *European journal of immunology* 26(1):28-33.
39. Akashi K, Kondo M, von Freeden-Jeffry U, Murray R, & Weissman IL (1997) Bcl-2 rescues T lymphopoiesis in interleukin-7 receptor-deficient mice. *Cell* 89(7):1033-1041.
40. Maraskovsky E*, et al.* (1997) Bcl-2 can rescue T lymphocyte development in interleukin-7 receptor-deficient mice but not in mutant rag-1-/- mice. *Cell* 89(7):1011-1019.
41. Beaudin AE, Boyer SW, & Forsberg EC (2014) Flk2/Flt3 promotes both myeloid and lymphoid development by expanding non-self-renewing multipotent hematopoietic progenitor cells. *Experimental hematology* 42(3):218-229 e214.
42. Dolence JJ, Gwin KA, Shapiro MB, & Medina KL (2014) Flt3 signaling regulates the proliferation, survival, and maintenance of multipotent hematopoietic progenitors that generate B cell precursors. *Experimental hematology* 42(5):380-393 e383.
43. Carvalho TL, Mota-Santos T, Cumano A, Demengeot J, & Vieira P (2001) Arrested B lymphopoiesis and persistence of activated B cells in adult interleukin 7(-/)- mice. *The Journal of experimental medicine* 194(8):1141-1150.
44. Jensen CT*, et al.* (2008) FLT3 ligand and not TSLP is the key regulator of IL-7-independent B-1 and B-2 B lymphopoiesis. *Blood* 112(6):2297-2304.
45. Vosshenrich CA, Cumano A, Muller W, Di Santo JP, & Vieira P (2003) Thymic stromal-derived lymphopoietin distinguishes fetal from adult B cell development. *Nature immunology* 4(8):773-779.
46. Rothenberg EV (2014) Transcriptional control of early T and B cell developmental choices. *Annual review of immunology* 32:283-321.
47. Ceredig R, Rauch M, Balciunaite G, & Rolink AG (2006) Increasing Flt3L availability alters composition of a novel bone marrow lymphoid progenitor compartment. *Blood* 108(4):1216-1222.

FIGURES

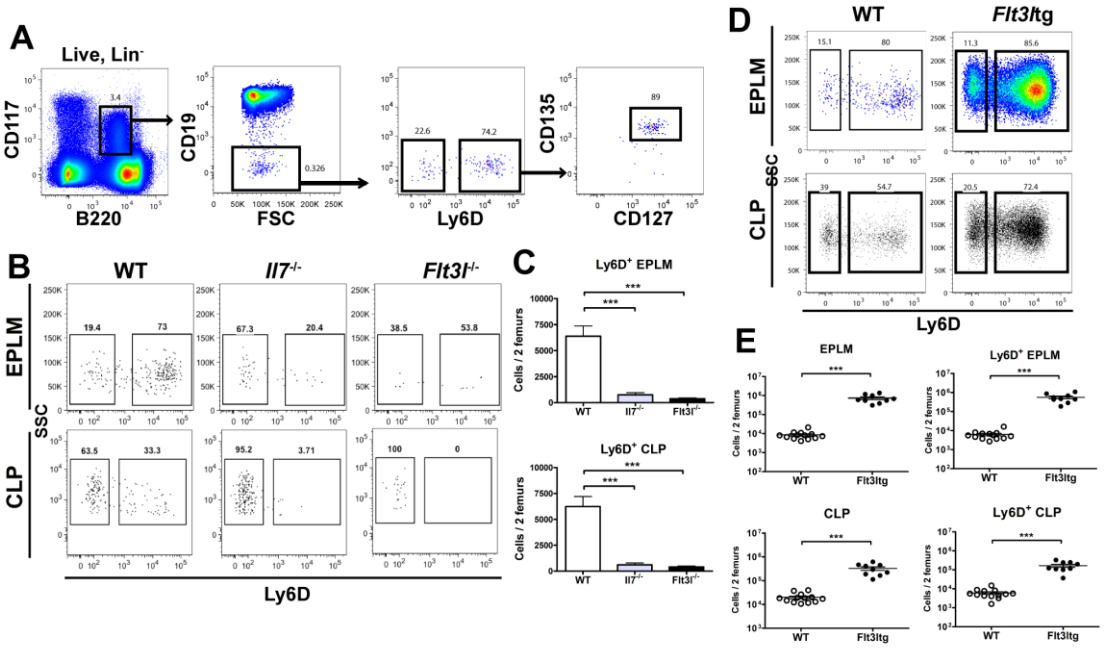


Figure 1. IL7 and FL are necessary for the generation of a normal Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ compartment. A. FACS plots showing the gating strategy used for identification of Ly6D$^+$ EPLM and their percentage of CD135 and CD127 expression. Lineage staining: SiglecH, CD115, CD11c, NK1.1, Gr-1. **B.** Representative FACS plots of EPLM (upper row) and CLP (lower row) from the bone marrow of WT, *Il7*$^{-/-}$ and *Flt3l*$^{-/-}$ mice. **C.** Absolute numbers of Ly6D$^+$ EPLM (upper graph) and CLP (lower graph) from the bone marrow of WT (n=13), *Il7*$^{-/-}$ (n=5) and *Flt3l*$^{-/-}$ (n=10) mice. **D.** Representative FACS plots of EPLM and CLP from WT and *Flt3Ltg* mice. **E.** Absolute numbers of total EPLM and CLP (left graphs) and Ly6D$^+$ EPLM and CLP (right graphs) from WT and *Flt3Ltg* mice.

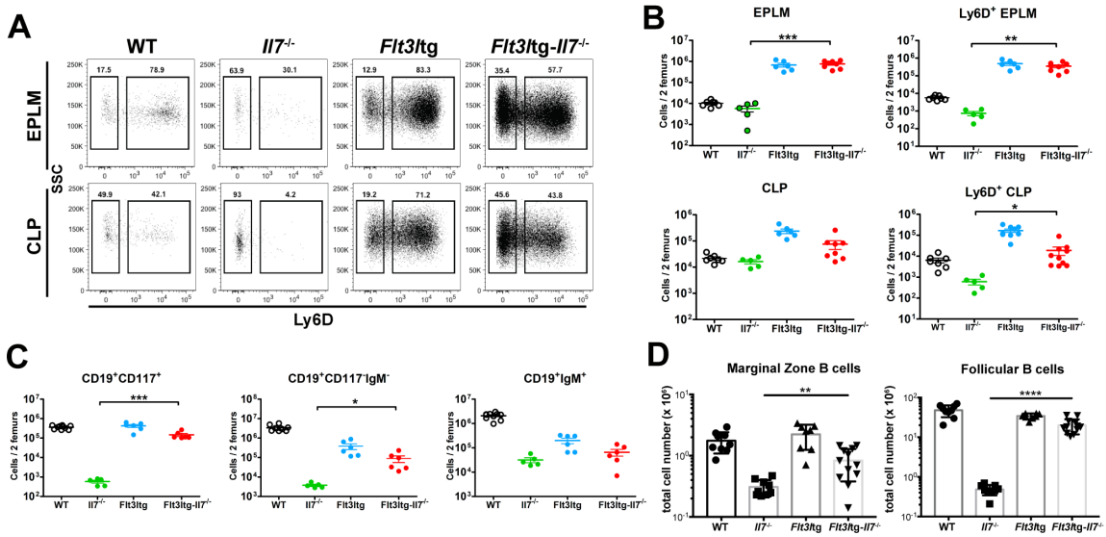


Figure 2. Increased *in vivo* FL levels rescue B cell generation in *Il7*$^{-/-}$ mice. A. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, *Il7*$^{-/-}$, *Flt3Ltg* and *Flt3Ltg-Il7*$^{-/-}$ mice. **B.** Numbers of EPLM (top left), CLP (bottom left), Ly6D^{+} EPLM (top right) and Ly6D^{+} CLP (bottom right) from the mouse genotypes indicated on the x-axes. For each mouse genotype mean±SEM is shown. **C.** Numbers of CD19^{+}CD117^{+} (top left), CD19^{+}CD117^{-}IgM^{-} (top right) and CD19^{+}IgM^{+} (bottom) bone marrow cells from the mice indicated on the x-axes. For each mouse genotype mean±SEM is shown. **D.** Numbers of CD19^{+}CD21^{high}CD23^{low} marginal zone (left) and CD19^{+}CD21^{+}CD23^{+} follicular (right) B cells in the spleens of WT or mutant mice, as indicated on the x-axes. For each mouse genotype mean±SD is shown.

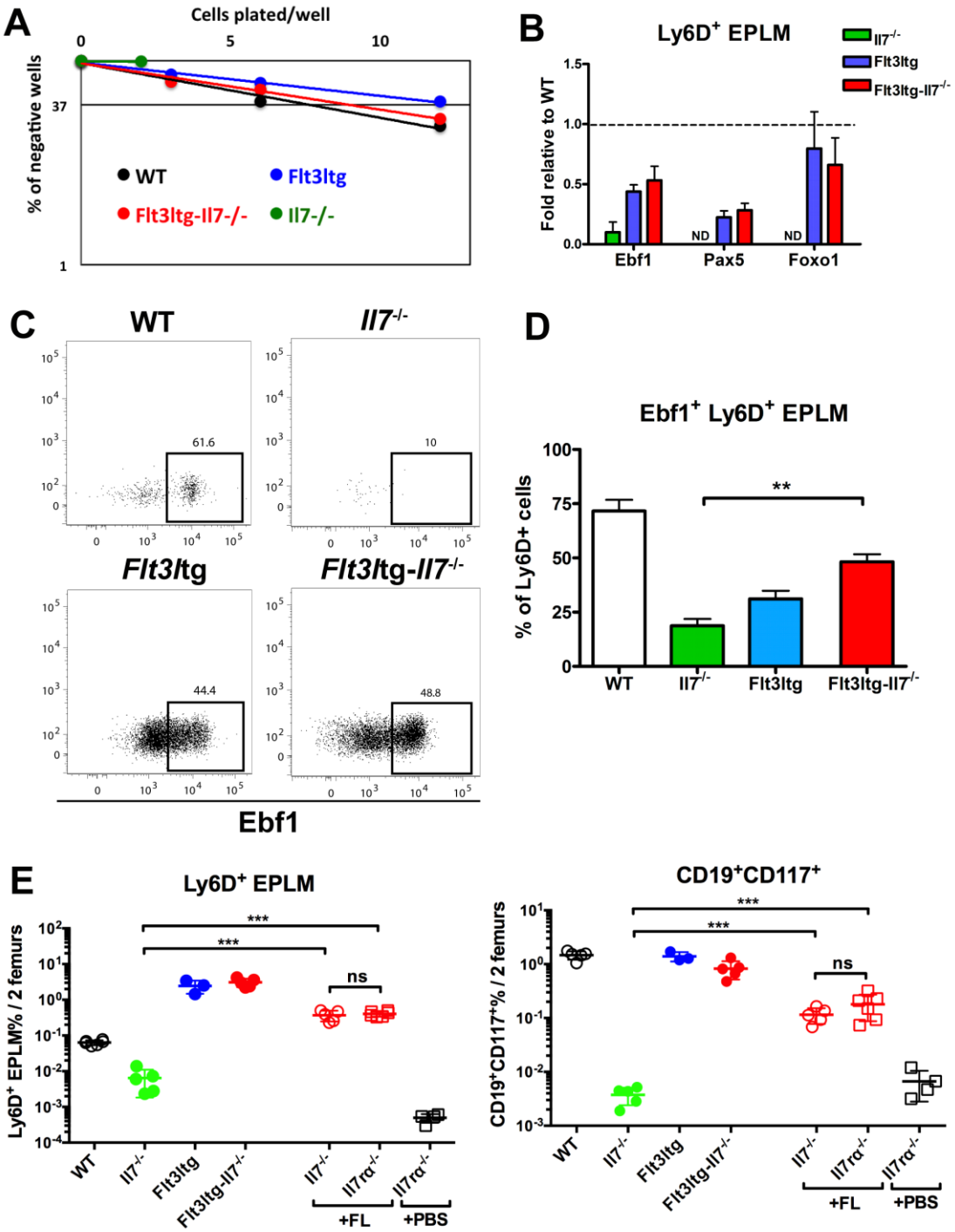


Figure 3. Increased *in vivo* FL rescues B cell commitment in the absence of IL7 and/or TSLP. A. *In vitro* limiting dilution analysis of Ly6D$^+$ EPLM B cell potential. Ly6D$^+$ EPLM were sorted from WT, *Il7*$^{-/-}$, *Flt3Ltg* and *Flt3Ltg-Il7*$^{-/-}$ mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. One representative out of four independent experiments is shown. **B.** Real-time quantitative PCR analysis showing expression of *Ebf1*, *Pax5* and *Foxo1* mRNAs in Ly6D$^+$ EPLM sorted from the indicated mouse genotypes. Bars show fold expression

relative to WT (set as 1). Error bars represent the SEM from 3-6 independent experiments. **C.** Representative FACS plots showing expression of Ebf1 protein within the $Ly6D^+$ EPLM of the indicated WT or mutant mice. **D.** Percentages of Ebf1-expressing $Ly6D^+$ EPLM from WT (n=7), *Il7*$^{-/-}$ (n=3), *Flt3Ltg* (n=11) and *Flt3Ltg-Il7*$^{-/-}$ (n=6) mice. Bars show mean±SEM. **E.** $Ly6D^+$ EPLM (left) and $CD19^+CD117^+$ (right) numbers from WT (n=5), *Il7*$^{-/-}$ (n=5), *Flt3Ltg* (n=3), *Flt3Ltg-Il7*$^{-/-}$ (n=5) mice, as well as from *Il7*$^{-/-}$ (n=5) and *Il7rα*$^{-/-}$ (n=6) mice injected intra-peritoneally with 10 daily doses of 10 µg FL each (indicated as +FL) or PBS (+PBS, n=4). Shown is the mean±SEM.

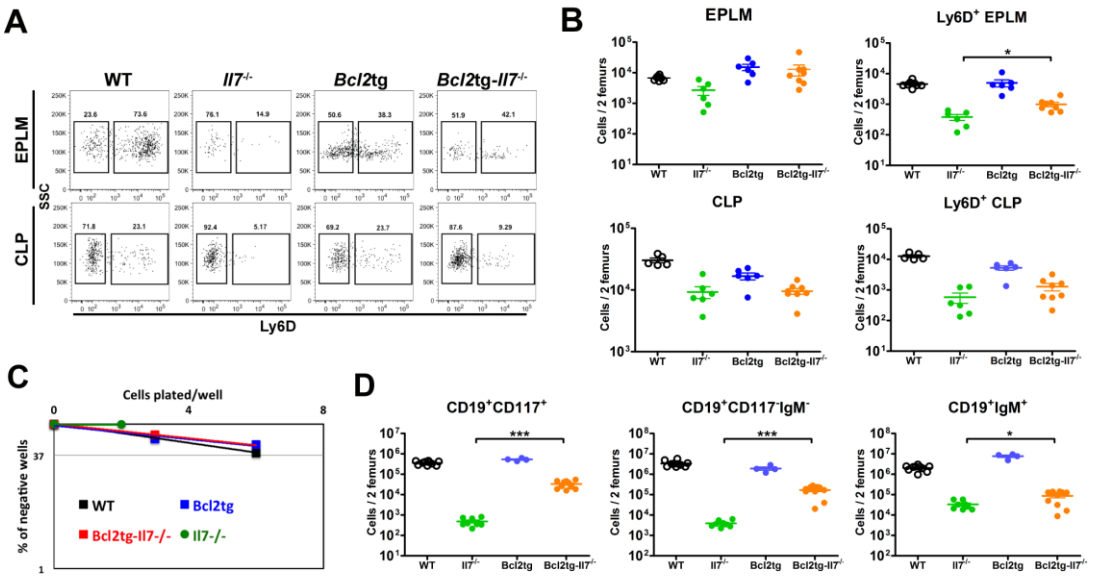


Figure 4. *Bcl2* over-expression partially rescues B cell commitment in *Il7*$^{-/-}$ mice. A. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, *Il7*$^{-/-}$, *Bcl2*tg and *Bcl2*tg-*Il7*$^{-/-}$ mice. **B.** Numbers of EPLM (top left), CLP (bottom left), Ly6D^{+} EPLM (top right) and Ly6D^{+} CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. **C.** *In vitro* limiting dilution analysis of Ly6D^{+} EPLM B cell potential. Ly6D^{+} EPLM were sorted from WT, *Il7*$^{-/-}$, *Bcl2*tg and *Bcl2*tg-*Il7*$^{-/-}$ mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. One representative out of three independent experiments is shown. **D.** Numbers of CD19^{+}CD117^{+} (top), CD19^{+}CD117^{-}IgM^{-} (middle) and CD19^{+}IgM^{+} (bottom) bone marrow cells from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown.

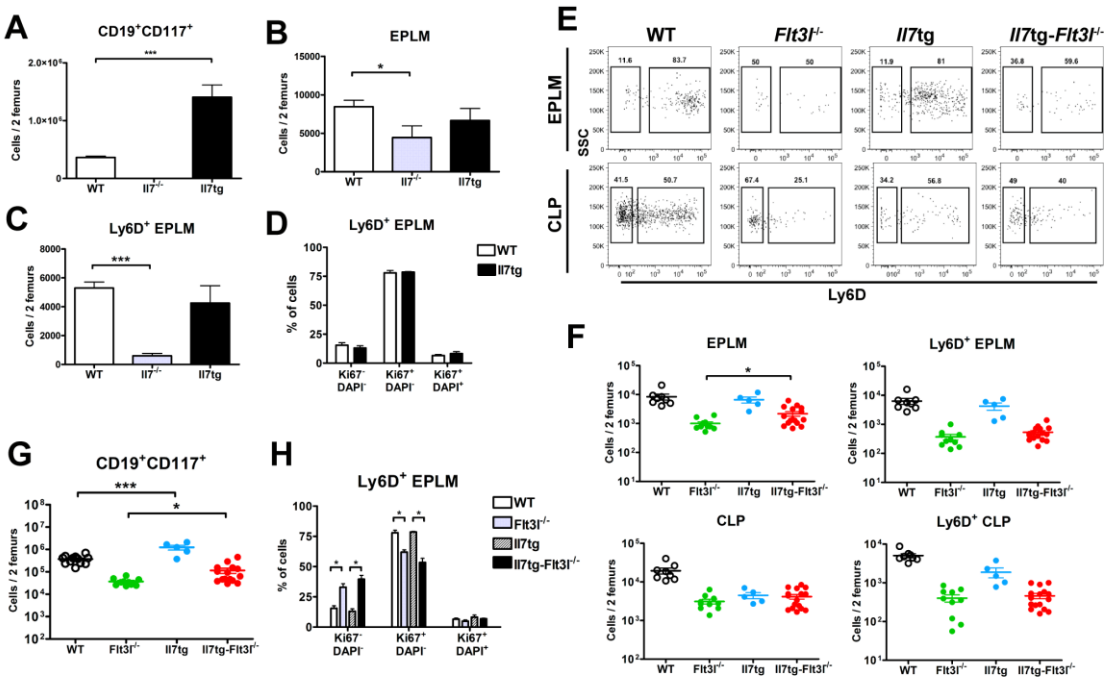


Figure 5. IL7 does not induce proliferation of Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ progenitors. A. CD19$^+$CD117$^+$ numbers in bone marrow of WT (n=10), *Il7*$^{-/-}$ (n=5) and *Il7*tg (n=8) mice. **B.** EPLM numbers in bone marrow of WT (n=14), *Il7*$^{-/-}$ (n=7) and *Il7*tg (n=5) mice. **C.** Ly6D$^+$ EPLM numbers in bone marrow of WT (n=14), *Il7*$^{-/-}$ (n=7) and *Il7*tg (n=5) mice. **D.** Cell cycle analysis of Ly6D$^+$ EPLM from WT (n=5) and *Il7*tg (n=2) mice. Graph shows percentages of Ki67$^-$DAPI$^-$, Ki67$^+$DAPI$^-$ and Ki67$^+$DAPI$^+$ Ly6D$^+$ EPLM. Bars in **A**, **B**, **C** and **D** show mean±SEM. **E.** Numbers of EPLM (top left), CLP (bottom left), Ly6D$^+$ EPLM (top right) and Ly6D$^+$ CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. **F.** Numbers of CD19$^+$CD117$^+$ bone marrow cells from WT and mutant mice, as indicated on the x-axis. For each mouse genotype mean±SEM is shown. **G.** Cell cycle analysis of Ly6D$^+$ EPLM from WT (n=5), *Flt3l*$^{-/-}$ (n=3), *Il7*tg (n=2) and *Il7*tg-*Flt3l*$^{-/-}$ (n=3) mice. Graph shows percentages of Ki67$^-$DAPI$^-$, Ki67$^+$DAPI$^-$ and Ki67$^+$DAPI$^+$ Ly6D$^+$ EPLM. Bars show mean±SEM.

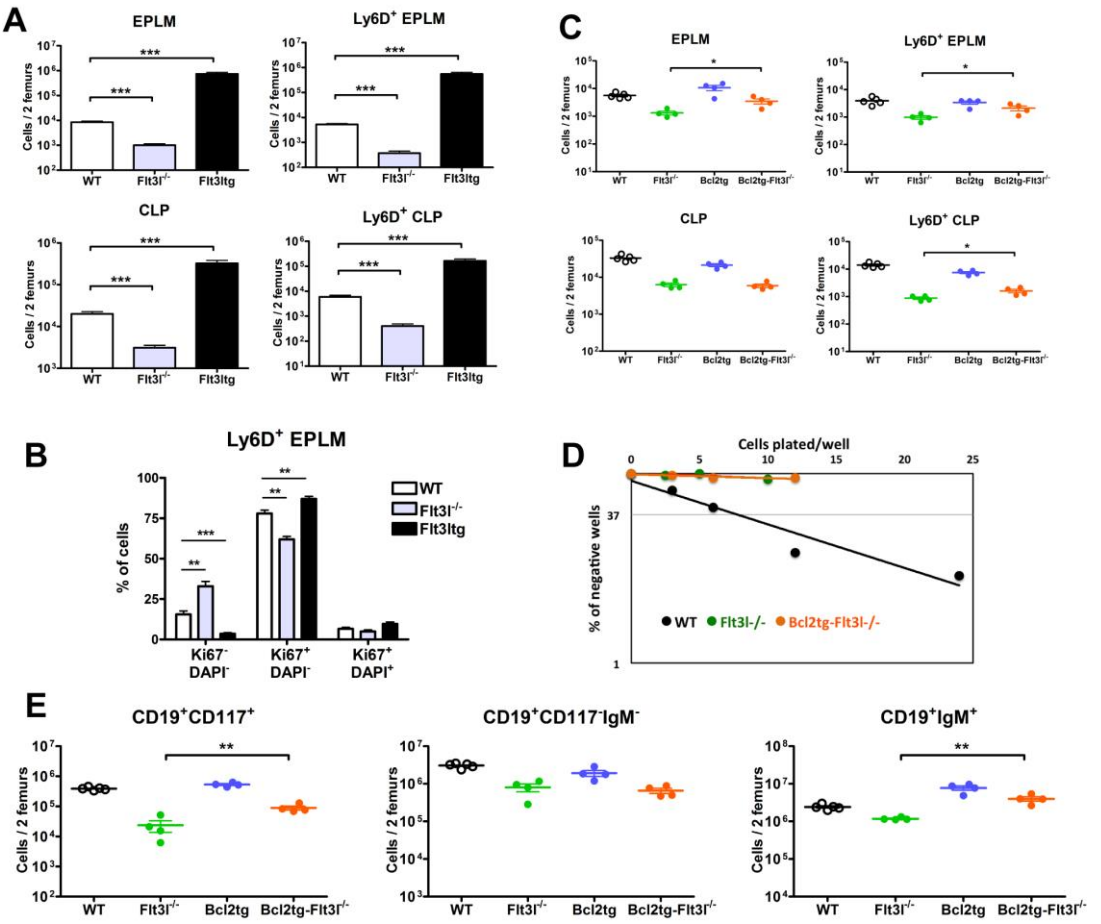


Figure 6. FL promotes proliferation but not survival of Ly6D$^+$CD135$^+$CD127$^+$CD19$^-$ progenitors. A. Numbers of EPLM (top left), CLP (bottom left), Ly6D$^+$ EPLM (top right) and Ly6D$^+$ CLP (bottom right) from WT (n=14), *Flt3l*$^{-/-}$ (n=10) and *Flt3Ltg* (n=9) mice. Bars show mean±SEM. **B.** Cell cycle analysis of Ly6D$^+$ EPLM from WT (n=5), *Flt3l*$^{-/-}$ (n=3) and *Flt3Ltg* (n=9) mice. Graph shows percentages of Ki67$^-$DAPI$^-$, Ki67$^+$DAPI$^-$ and Ki67$^+$DAPI$^+$ Ly6D$^+$ EPLM. Bars show mean±SEM. **C.** Numbers of EPLM (top left), CLP (bottom left), Ly6D$^+$ EPLM (top right) and Ly6D$^+$ CLP (bottom right) from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown. **D.** *In vitro* limiting dilution analysis of Ly6D$^+$ EPLM B cell potential. Ly6D$^+$ EPLM were sorted from WT, *Flt3l*$^{-/-}$ and *Bcl2*tg-*Flt3l*$^{-/-}$ mice and plated at the indicated concentrations on OP9 stromal cells together with IL7. **E.** Numbers of CD19$^+$CD117$^+$ (left), CD19$^+$CD117$^-$IgM$^-$ (middle) and CD19$^+$IgM$^+$ (right) bone marrow cells from WT and mutant mice, as indicated on the x-axes. For each mouse genotype mean±SEM is shown.

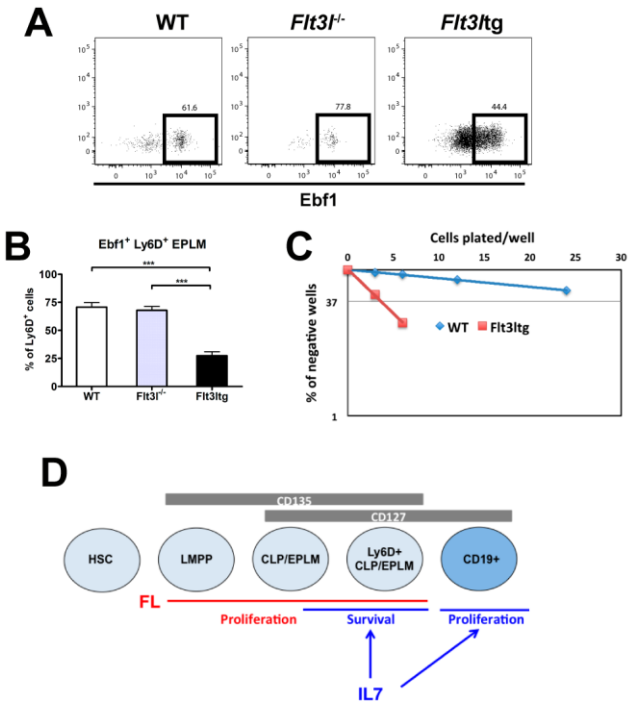


Figure 7. FL does not instruct *Ebf1* expression and B cell commitment. A. Representative FACS plots showing expression of Ebf1 protein within the Ly6D$^+$ EPLM of WT, *Flt3l*$^{-/-}$ and *Flt3Ltg* mice. **B.** Percentages of Ebf1-expressing Ly6D$^+$ EPLM from WT (n=7), *Flt3l*$^{-/-}$ (n=5) and *Flt3Ltg* (n=12) mice. Bars show mean±SEM. **C.** *In vitro* limiting dilution analysis of Ly6D$^+$ EPLM T cell potential. Ly6D$^+$ EPLM were sorted from WT and *Flt3Ltg* mice and plated at the indicated concentrations on OP9DL1 stromal cells together with IL7. One representative of four independent experiments is shown. **D.** Schematic model for the permissive role of IL7 and FL acting on hematopoietic progenitors and CD19$^+$ committed B-cell precursors. HSC: Hematopoietic Stem Cell; LMPP: Lymphoid-primed Multi-Potent Progenitor; CLP: Common Lymphoid Progenitor; EPLM: Early Progenitor with Lymphoid and Myeloid potential.

SUPPORTING INFORMATION

Supporting Materials and Methods

Antibodies

The following antibodies were used for flow cytometry (from BD Pharmingen, eBioscience, BioLegend, or produced in house): anti-B220 (RA3-6B2), anti-CD117 (2B8), anti-CD19 (1D3), anti-NK1.1 (PK136), anti-SiglecH (551), anti-CD11c (HL3), anti-CD115 (AFS98), anti-Ly6D (49-H4), anti-CD127 (SB/199), anti-CD135 (A2F10), anti-Sca1 (D7), anti-IgM (M41), anti-CD21 (7G6), anti-CD23 (B3B4), anti-CD4 (GK1.5), anti-CD8 (53.6.7), anti-TCRβ (H57).

Quantitative real-time PCR

The primers used were: *Ebf1:* Ebf1-F: 5'-CAGGAAACCCACGTGACAT-3'; Ebf1-R: 5'-CCACGTTGACTGTGGTAGACA-3', *Pax5:* Pax5-F: 5'ACGCTGACAGGGATGGTG-3'; Pax5-R: 5'-GGGGAACCTCCAAGAATCAT-3', *Foxo1:* Foxo1-F: 5'-AGTGGATGGTGAAGAGCGT-3', Foxo1-R: 5'-GAAGGGACAGATTGTGGCG-3', *Actin:* Actin-F: 5'-CTGTCGAGTCGCGTCCACC-3', Actin-R: 5'-CGCAGCGATATCGTCATCCA-3'.

Supporting Figures

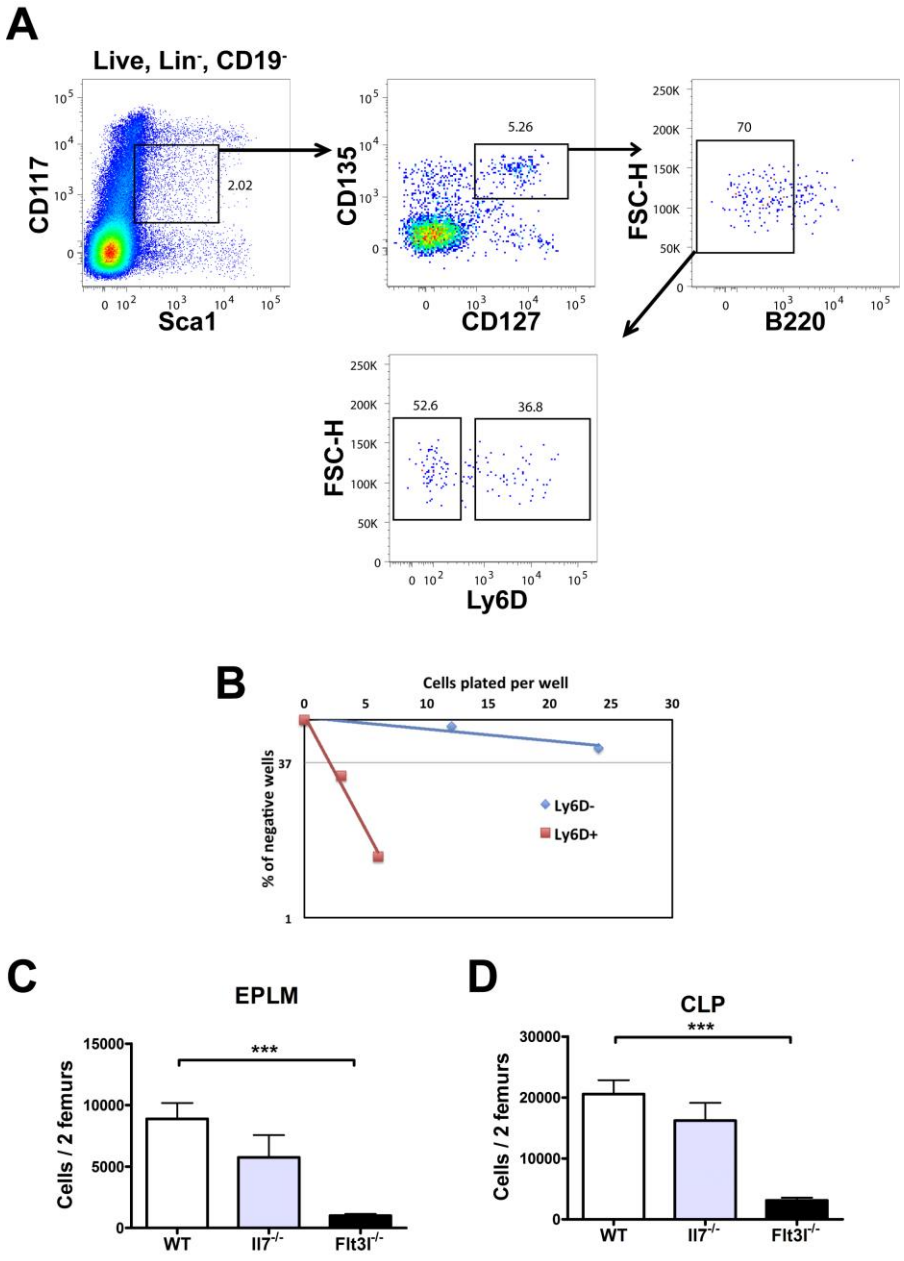


Fig. S1. A. CLP FACS staining in WT mice. FACS plots showing the gating strategy used for the identification of Ly6D^{+} CLP. Lineage staining: SiglecH, CD115, CD11c, NK1.1, Gr-1. **B.** *In vitro* limiting dilution analysis of Ly6D^{+} and Ly6D^{-} EPLM B cell potential. Cells were sorted as shown in Figure 1A and plated at the indicated concentrations on OP9 stromal cells together with IL7 A representative of 3 independent experiments is shown**. C, D**. Numbers of EPLM (**C**) and CLP (**D**) progenitors in WT (n=13), *Il7*$^{-/-}$ (n=5) and *Flt3l*$^{-/-}$ (n=10) mice. EPLM were stained as shown in Figure 1A and CLP as shown in **A**. Student's t test. ***P ≤ 0.001. Bars show mean±SEM.

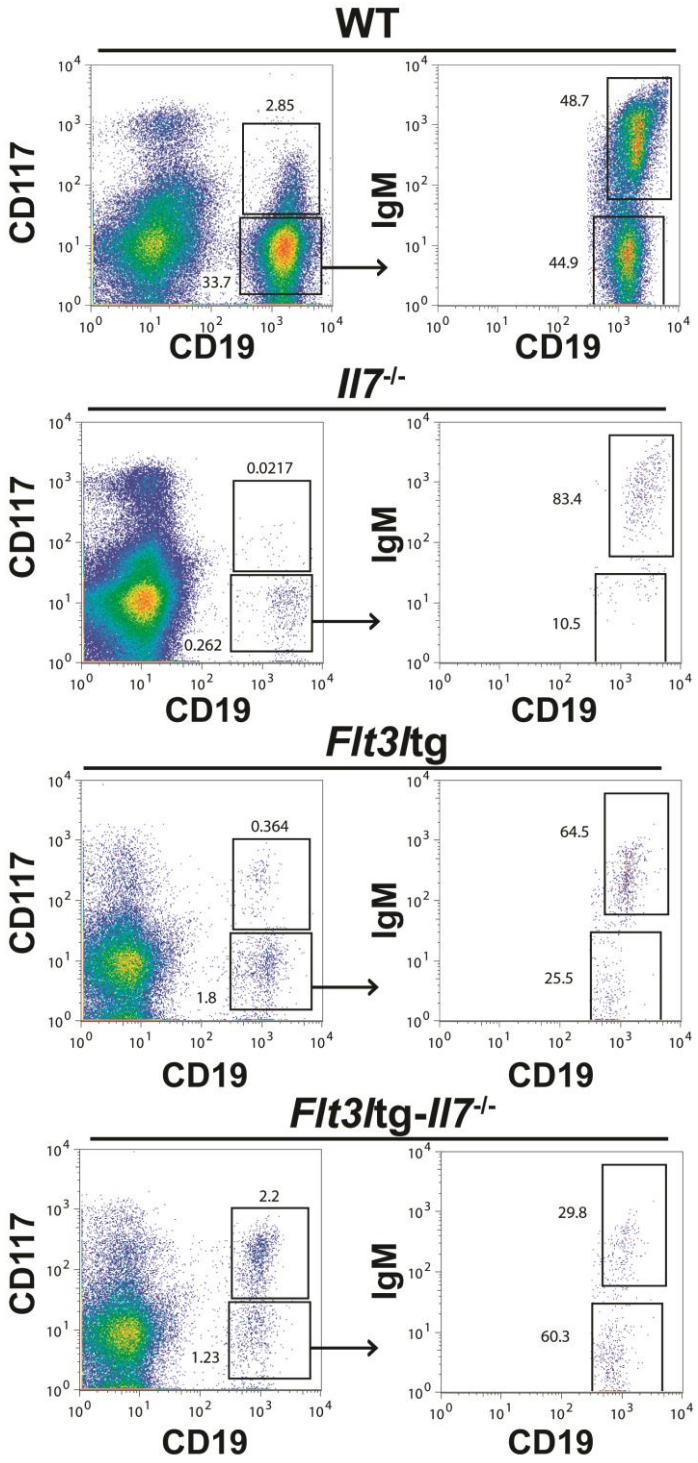


Fig. S2. Rescue of CD19^{+} bone marrow B cell progenitors in *Flt3Ltg-Il7*$^{-/-}$ mice. Figure shows representative FACS plots for the identification of CD19^{+}CD117^{+}, CD19^{+}CD117^{-} IgM^{-} and CD19^{+}IgM^{+} bone marrow cells.

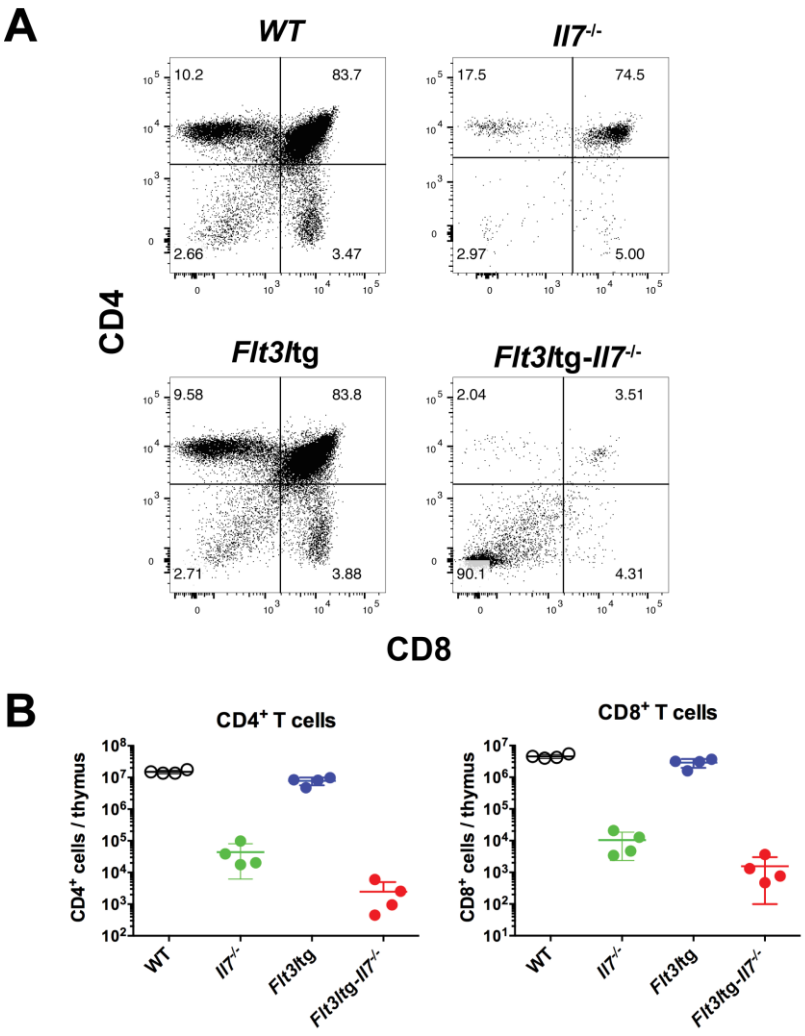


Fig. S3. Thymic T cell development in *Flt3Ltg-Il7*$^{-/-}$ mice. A. Representative FACS plots showing CD4/CD8 thymocyte staining from 6-8 week old WT, *Il7*$^{-/-}$, *Flt3Ltg* and *Flt3Ltg-Il7*$^{-/-}$ mice (n=4 for each group). **B.** Total numbers of CD4^{+} (left panel) and CD8^{+} (right panel) single-positive thymocytes from the mouse genotypes indicated on the x-axes.

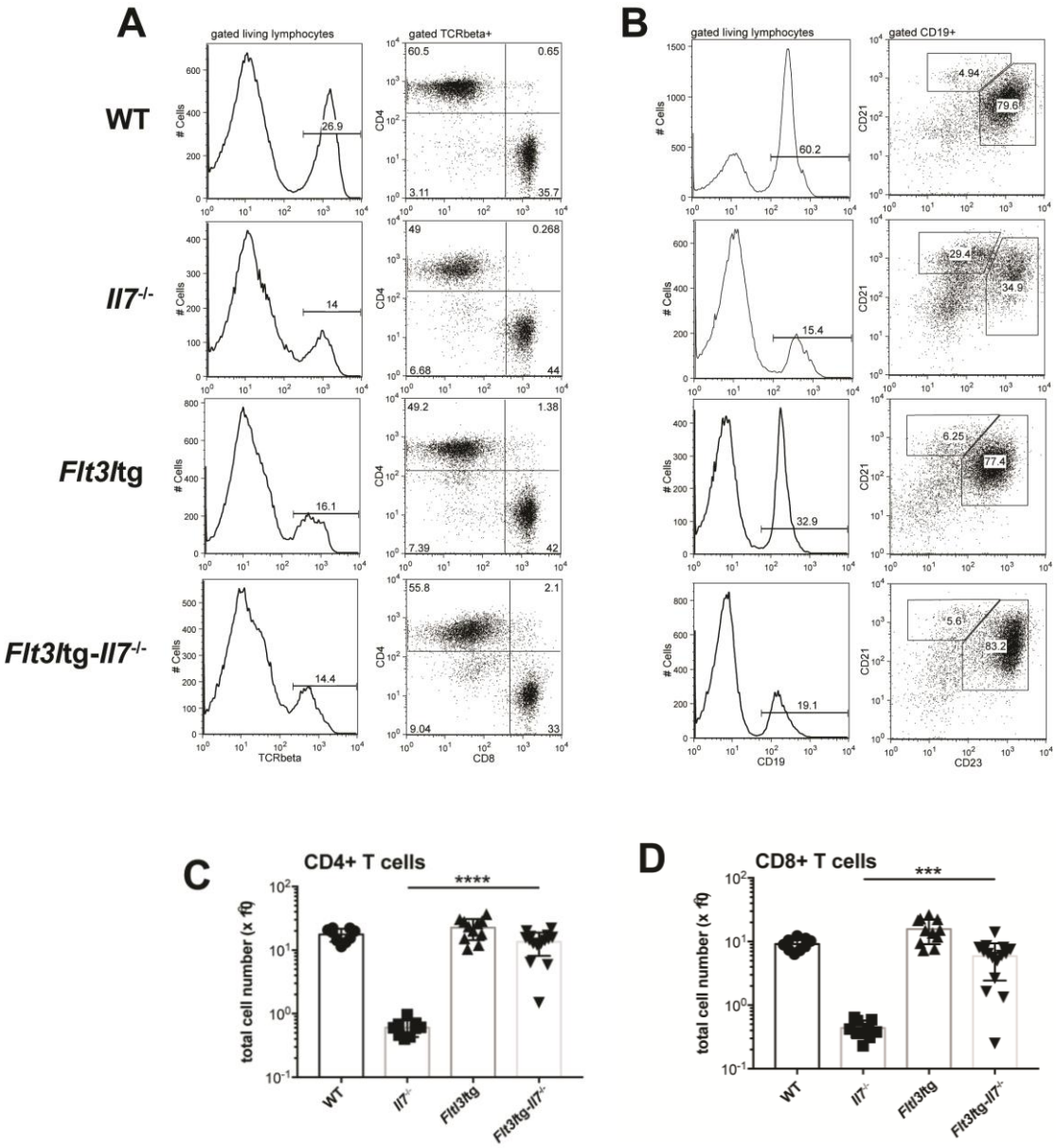


Fig. S4. A. Representative FACS plots illustrating T cells in the spleens of WT (first row), *Il7*$^{-/-}$ (second row), *Flt3Ltg* (third row), and *Flt3Ltg-Il7*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes TCRβ^{+} cells are further sub-grouped in CD4 and CD8 positive T cells. **B**. Representative FACS plots illustrating B cells in the spleens of WT (first row), *Il7*$^{-/-}$ (second row), *Flt3Ltg* (third row), and *Flt3Ltg-Il7*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes CD19^{+} cells are further sub-grouped in CD21^{high}CD23^{low} marginal zone B cells and CD21^{+}CD23^{+} follicular B cells. **C** and **D**. Numbers of splenic CD4^{+} (C) and CD8^{+} (D) T cells, stained as shown in A, from WT and mutant mice as indicated on the x-axes. ***P ≤ 0.001, ****P ≤ 0.0001. Student's t test; n = 9-15. Data shown above are mean ± SD.

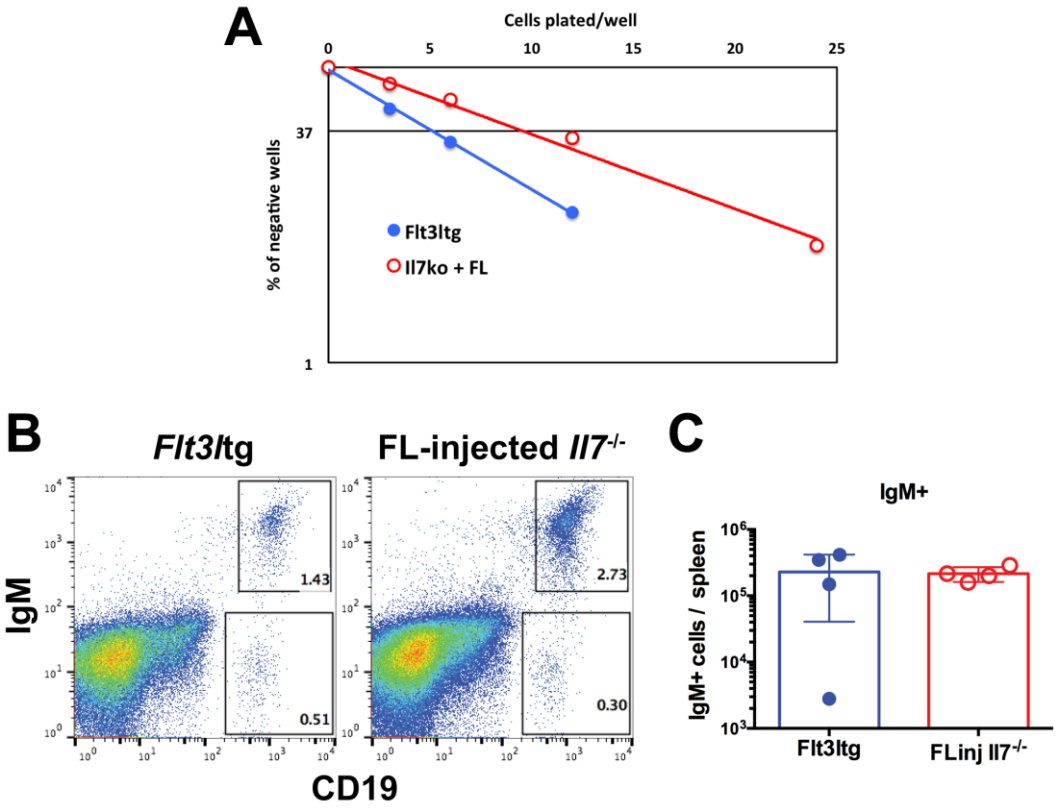


Fig. S5. B cell potential of Ly6D$^+$ EPLM cells from *Il7*$^{-/-}$ mice injected with FL. *Il7*$^{-/-}$ mice were injected with FL (10 daily doses of 10µg for each mouse) and Ly6D$^+$ EPLM were sorted from their bone marrows one day after the last injection **A**. *In vitro* limiting dilution analysis of the B cell potential of FL-injected *Il7*$^{-/-}$ Ly6D$^+$ EPLM. Cells were plated at the indicated concentrations on OP9 stromal cells plus IL7. *Flt3Ltg* Ly6D$^+$ EPLM were used as positive controls. **B, C.** *In vivo* B cell potential of FL-injected *Il7*$^{-/-}$ Ly6D$^+$ EPLM. Five thousand Ly6D$^+$ EPLM from FL-injected *Il7*$^{-/-}$ or *Flt3Ltg* mice were intravenously injected into sub-lethally irradiated *Rag2*$^{-/-}$ mice. Four weeks after cell transfer spleens were analyzed for expression of CD19 and IgM. **B.** Respresentative FACS plots of recipient spleens. **C.** Numbers of CD19$^+$IgM$^+$ B cells harvested from the analyzed spleens (n=4 mice for each group).

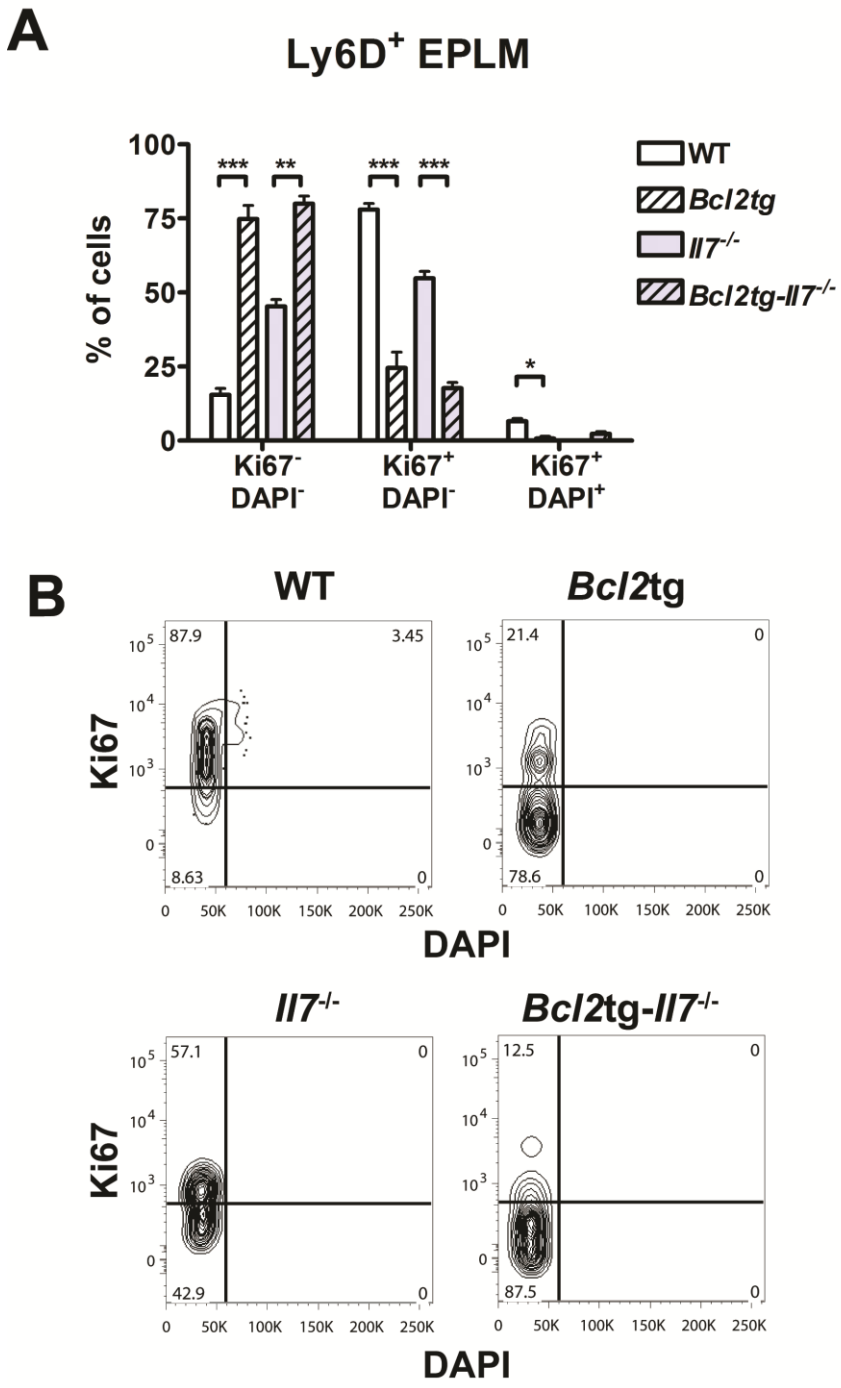


Fig. S6. Quiescent state of *Bcl2*-rescued cells *in vivo*. A. Cell cycle analysis of Ly6D$^+$ EPLM from WT (n=5), *Bcl2*tg, (n=2), *Il7*$^{-/-}$ (n=2) and *Bcl2*tg-*Il7*$^{-/-}$ (n=4) mice. Graph shows percentages of Ki67$^-$DAPI$^-$, Ki67$^+$DAPI$^-$ and Ki67$^+$DAPI$^+$ Ly6D$^+$ EPLM. *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001. Student's t test. Bars show mean±SEM. **B**. Representative Ki67/DAPI FACS plots of the Ly6D$^+$ EPLM cell cycle analysis collectively presented in **A.**

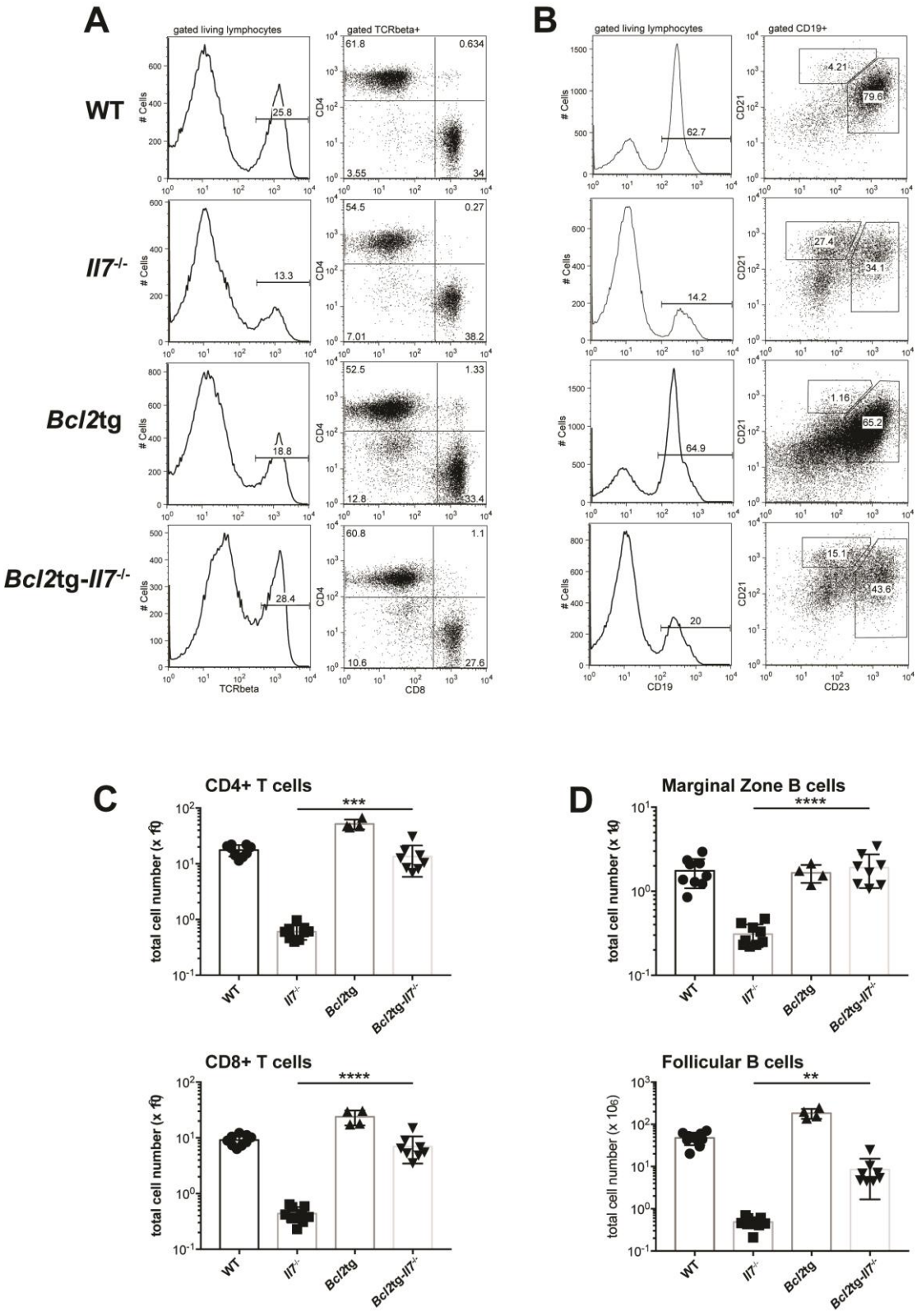


Fig. S7. Bcl2-mediated rescue of splenic T and B cells in the absence of IL7. A. Representative FACS plots illustrating T cells in the spleens of WT (first row), *Il7*$^{-/-}$ (second row), *Bcl2*tg (third row), and *Bcl2*tg-*Il7*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes TCRβ^{+} cells are further sub-grouped in CD4 and CD8 positive T

cells. **B**. Representative FACS plots illustrating B cells in the spleens of WT (first row), *Il7*$^{-/-}$ (second row), *Bcl2*tg (third row), and *Bcl2*tg-*Il7*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes CD19^{+} cells are further sub-grouped in CD21^{high}CD23^{low} marginal zone B cells and CD21^{+}CD23^{+} follicular B cells. **C**. Numbers of splenic CD4^{+} (top) and CD8^{+} (bottom) T cells, stained as shown in **A**, from WT and mutant mice as indicated on the x-axes. **D**. Numbers of splenic marginal zone (top) and follicular (bottom) B cells, stained as shown in **B**, from WT and mutant mice as indicated on the x-axes. **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Student's t test; n = 4-9. Data shown above are mean ± SD.

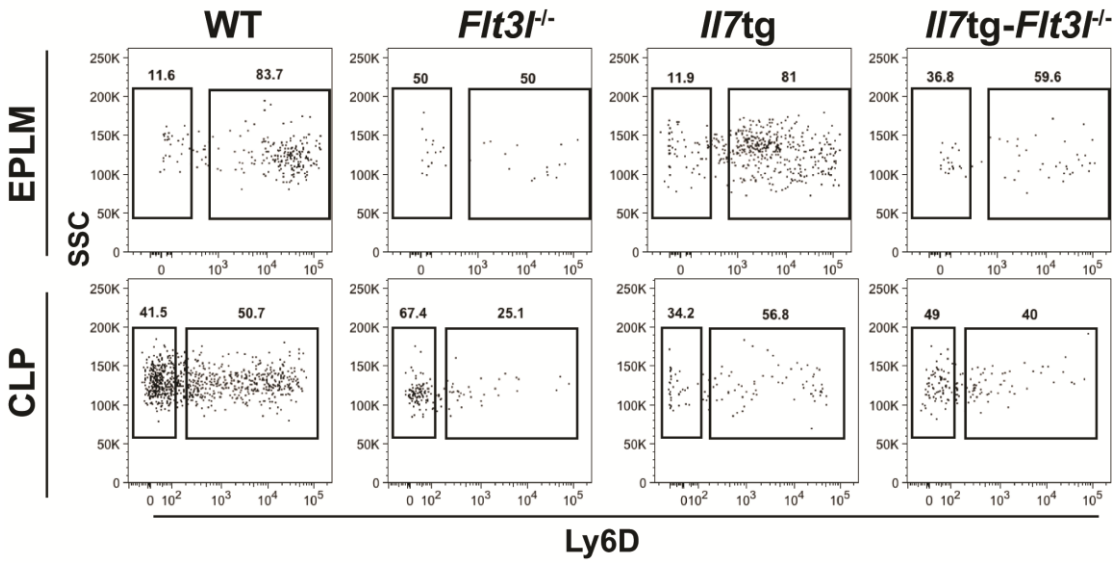


Fig. S8. Representative FACS plots of EPLM (upper panel) and CLP (lower panel) from WT, *Flt3l*$^{-/-}$, *Il7*tg and *Il7*tg-*Flt3l*$^{-/-}$ mice.

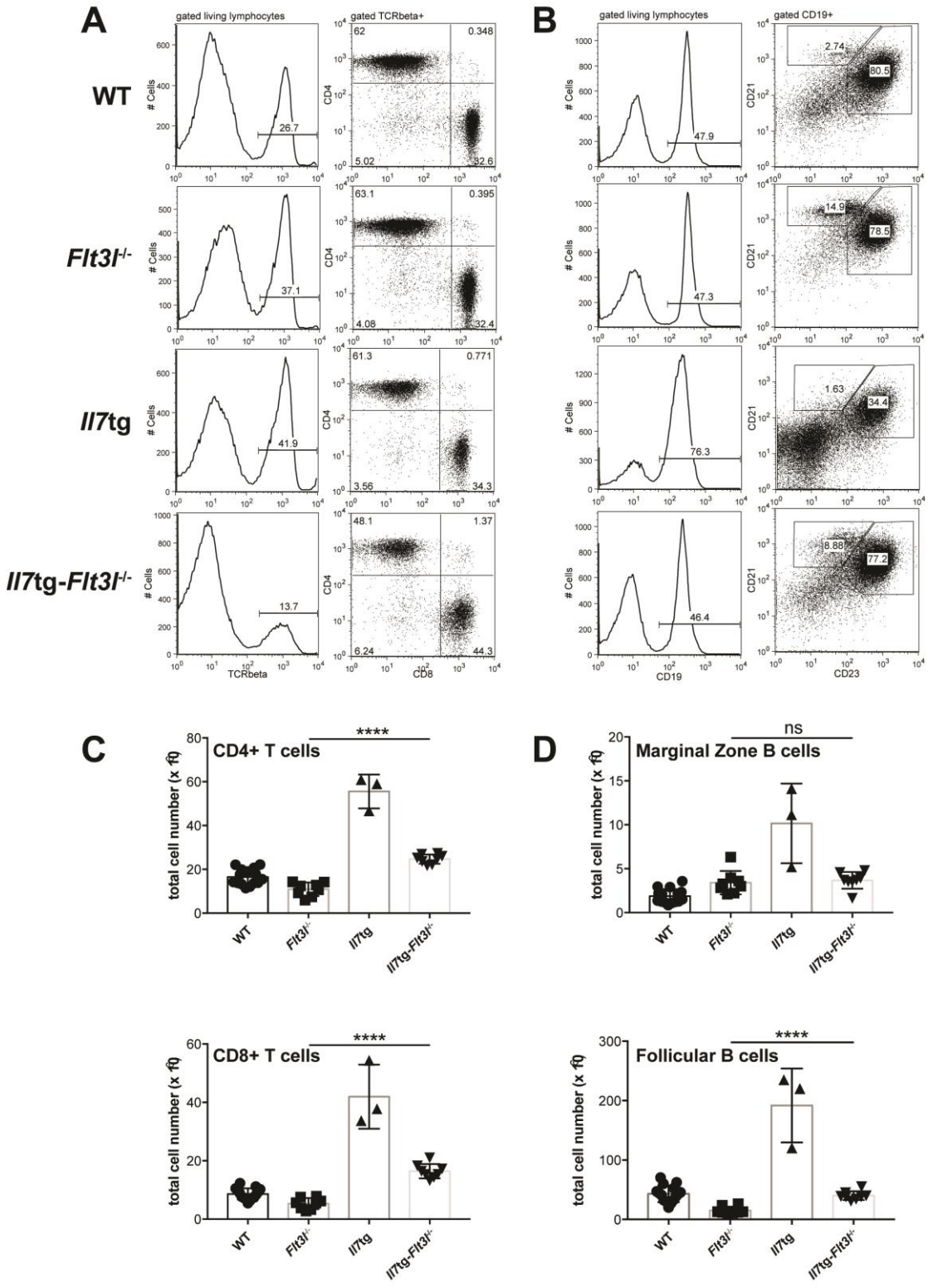


Fig. S9. Effect of IL7 over-expression on WT and *Flt3l*$^{-/-}$ splenic T and B cells. A. Representative FACS plots illustrating T cells in the spleens of WT (first row), *Flt3l*$^{-/-}$ (second row), *Il7*tg (third row), and *Il7*tg-*Flt3l*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes TCRβ^{+} cells are further sub-grouped in CD4 and CD8 positive T cells. **B**. Representative FACS plots illustrating B cells in the spleens of WT (first row),

Flt3l$^{-/-}$ (second row), *Il7*tg (third row), and *Il7*tg-*Flt3l*$^{-/-}$ (fourth row) mice. After gating on living lymphocytes CD19^{+} cells are further sub-grouped in CD21^{high}CD23^{low} marginal zone B cells and CD21^{+}CD23^{+} follicular B cells. **C**. Numbers of splenic CD4^{+} (top) and CD8^{+} (bottom) T cells, stained as shown in **A**, from WT and mutant mice as indicated on the x-axes. **D**. Numbers of splenic marginal zone (top) and follicular (bottom) B cells, stained as shown in **B**, from WT and mutant mice as indicated on the x-axes. ns not significant or $P > 0.05$, ****$P \leq 0.0001$. Student's t test; n = 3-15. Data shown above are mean ± SD.

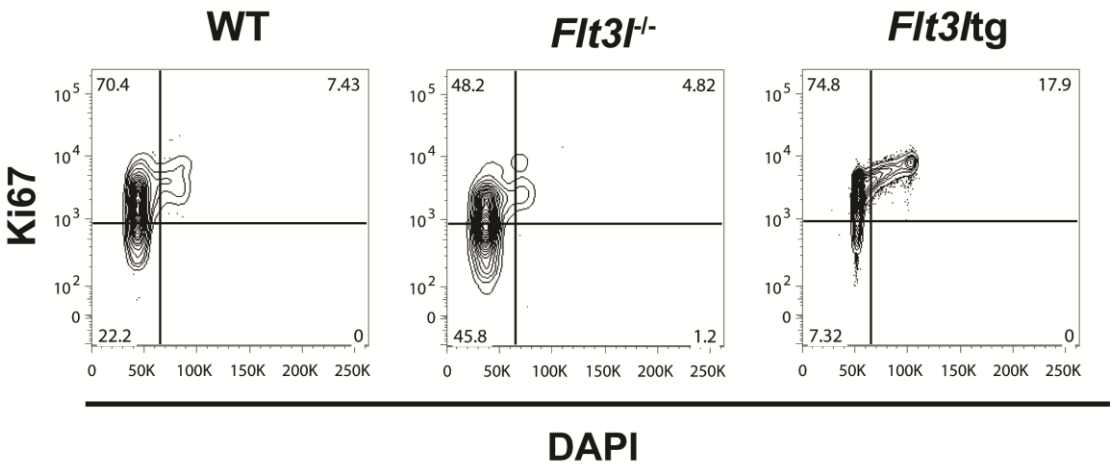


Fig. S10. Effect of *in vivo* FL levels on Ly6D$^+$ EPLM cell cycle. Representative Ki67/DAPI FACS plots of the Ly6D$^+$ EPLM cell cycle analysis collectively presented in Figure 6B.

9.2 Paper 2:

Reconstitution of a functional B-cell compartment in immunodeficient mice with pro-B cells propagated with or without stromal cells

Lilly von Muenchow[1], Panagiotis Tsapogas[1], Llucia Albertí-Servera[1], Giuseppina Capoferri[1], Marianne Doelz[1], Hannie Rolink[1], Nabil Bosco[1, 2], Rhodri Ceredig[3] and Antonius G. Rolink[1]

[1] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[2] Present address: Nestlé Research Center, Lausanne, Switzerland

[3] Department of Biosciences, University of Galway, Galway, Ireland

Correspondence: Prof. Antonius G. Rolink; e-mail: antonius.rolink@unibas.ch

Acknowledgements

A.G.R. is holder of the chair in immunology endowed by L. Hoffmann – La Roche Ltd, Basel. This study was supported by the Swiss National Science Foundation and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement number 315902. R.C. was supported by Science Foundation Ireland under grant numbers SFI09/SRC/B1794 and SFI07/SK/B1233b.

Conflict of interest

The authors declare no commercial or financial conflict of interest.

ABSTRACT

Fetal liver (FL) and bone marrow (BM) derived pro-B cells were propagated long-term in stromal cell free cultures supplemented with interleukin-7 (IL-7), stem cell factor and FLT3 ligand. Within a week, most cells expressed surface CD19, CD79A, λ5 and VpreB antigens and had rearranged immunoglobulin D-J heavy chain genes. Both FL and BM pro-B cells reconstituted the B-cell compartments of immuno-incompetent Rag2-deficient mice with FL pro-B generating follicular, marginal zone (MZB) and B1a B cells, but BM pro-B cells mainly MZB. Reconstituted mice generated significant IgM and IgG antibodies to a type II T-independent antigen with FL pro-B cell reconstituted mice generating surprisingly high IgG_1 titers. Finally, we show for the first time that mice reconstituted with mixtures of stromal cell free *in vitro* propagated pro-B and pro-T cells mounted a T-cell dependent antibody response. This novel stromal cell free culture system facilitates our understanding of B-cell development and might be applied clinically.

INTRODUCTION

In mammals, B cells develop from hematopoietic stem cells (HSCs), antenatally from the fetal liver (FL) and postnatally from the bone marrow (BM). Various stages of B-cell development are distinguishable by combinations of cell surface and intracellular markers, cell cycle profile and rearrangement status of IgH and IgL genes [1-4]. B-cell commitment is determined by the transcription factor Pax5 [5-7] one of whose target genes is CD19 [8]. The earliest B-cell committed precursor, or pro-B cell, proliferates rapidly, is $CD19^+$ $CD117^+$ [1, 9] and has its Ig_H D–J genes rearranged [4] and are absent in Pax5-deficient [10], *Il7* or *Il7R* gene deleted mice [11, 12]. The receptor tyrosine kinases CD117 and CD135 and their corresponding ligands, SCF and FLT3L, are important for early B-cell development [13-17]. Blocking SCF binding with an anti-CD117 antibody inhibits pro-B cell proliferation in IL-7-containing stromal cell cultures [18] and CD135 or FLT3L-deficient mice have a dramatically reduced BM pro-B-cell compartment [19]. Thus Pax5 and cytokines are important for early B-cell development.

We previously showed that FL-derived pro-B cells could be grown long-term on stromal cells and IL-7 [20]. Here we present for the first time that this is possible for BM-derived pro-B cells. We have described an early (E) progenitor (P) with lymphoid (L) and myeloid (M) developmental potential, called EPLM, in the BM [21]. EPLM are $B220^+$, $CD117^+$ but $CD19^-$ $NK1.1^-$. When cultured either on OP9 stromal cells plus IL-7 or without stromal cells but with IL-7, SCF and FLT3L, EPLM from the FL or BM can be propagated long-term and differentiate into $CD19^+$ pro-B cells. Moreover, upon *in vivo* transplantation into immunodeficient *Rag2* gene deficient hosts, these pro-B cells reconstitute a functional B-cell compartment. Finally, when these mice are reconstituted with a mixture of *in vitro* stromal cell free propagated pro-B and pro-T cells [22], a small, but functional adaptive immune system is generated.

RESULTS

Growth of BM pro-B cells long-term

Using stromal cell-based cultures, only FL-derived pro-B cells could be grown long-term [20]. We recently identified a $B220^{+}CD117^{+}CD19^{-}NK1.1^{-}$ BM progenitor having lymphoid and myeloid developmental potential, called EPLM [21]; now we tested their capacity to generate long-term growing pro-B cells. Thus, $B220^{+}CD117^{+}CD19^{-}NK1.1^{-}$ ($CD19^{-}$) and $B220^{+}CD117^{+}CD19^{+}NK1.1^{-}$ ($CD19^{+}$) cells from FL or BM were sorted (Fig.1A) and plated on OP9 stromal cells plus IL-7. After 6 days, cells were harvested every 3 - 4 days and re-plated on fresh stromal cells and IL-7. FACS staining 6-7 days after initiation of cultures showed in all cases that the cultured cells were >90% positive for CD19 expression and this expression was further increased to almost 100% and retained throughout the culture period (data not shown and Figure 1D). $CD19^{-}$ (EPLM) cells grew continuously with doubling times of about 30hrs, whereas $CD19^{+}$ cells proliferated for the first 7 days and then died (Fig 1B). RNAseq analysis on $CD19^{+}$ cells derived from $CD19^{-}$ EPLM grown for 12 days on OP9 plus IL-7 and freshly-isolated $CD19^{+}CD117^{+}$ cells similarly grown for 5 days revealed that only 83 genes, none of which were B-cell related, were ≥2 fold differentially expressed (Table SI). All 44 B cell related genes were similarly expressed (Table SII). Gene ontology analysis of the 83 genes did not identify genes that explained the growth difference observed. An additional RNAseq analysis of *ex vivo* isolated $CD19^{+}CD117^{+}$ cells and EPLM derived $CD19^{+}$ cells maintained on OP9 stromal cells in the presence of IL-7 for two weeks revealed 3449 differentially expressed genes (Table SIII). However, amongst them we could not identify particular candidates that might be responsible for the dissimilar growth capacity observed.

FL-derived EPLM grew better than BM-derived EPLM on OP9 with IL-7 (Fig.1C) with >95% of both becoming $CD19^{+}$ (Fig.1D) with ~25% BM and >60% FL-derived cells expressing CD117 (Fig.1D). RT-PCR analysis revealed that both expressed CD79a, CD79b, Igll1 (λ5) and Vpreb1 (Fig.1E) and both had undergone D_H-J_H rearrangements (Fig.1F). Thus, BM and FL EPLM cultured on OP9 plus IL-7 give rise to pro-B cells proliferating for more than three months.

Long-term propagation of BM and FL pro-B cells in stromal cell free cultures

FL pro-B cells propagated on stromal cells plus IL-7 reconstituted the B-cell compartment of immuno-deficient mice [20, 29, 30]. However, it is unlikely that pro-B cells derived from co-culture settings will ever be approved for therapeutic purposes. Therefore, we developed a stromal cell free culture system consisting of soluble IL-7, SCF and FLT3L. As shown in figure 2A and B, EPLM from FL and BM showed very robust growth under these stromal cell free conditions. Indeed, the growth rate of BM-derived EPLM was identical under stromal cell and stromal cell free conditions (Fig.2A). However, growth of FL EPLM was slightly slower without stromal-cell support (Fig.2B). Stromal cell free cultured EPLM became $CD19^+$ with ~35% BM and >75% FL-derived cells expressing CD117 (Fig.2C and D). Moreover, cells cultured without stroma expressed CD79a and b, Igll1 (λ5) and Vpreb1 (Fig.2E) and had their Ig_H chains D_H-J_H rearranged (Fig.2F). Thus, both BM and FL EPLM cultured with IL-7, SCF and FLT3L alone also give rise to long-term proliferating pro-B cells. Moreover, as shown for FL-derived pro-B cells [20], BM-derived pro-B cells expressing the anti-apoptotic Bcl2 transgene cultured with or without stromal-cell support efficiently differentiated into IgM^+ B cells upon IL-7 removal (Figure S1).

***In vivo* B-cell reconstitution by BM or FL EPLM-derived pro-B cells.**

To test whether *in vitro*-generated FL or BM-derived pro-B cells could reconstitute mice, 10^7 EPLM-derived pro-B from FL or BM of B6 CD45.1 mice cultured with or without stroma were transferred into sub-lethally irradiated CD45.2 B6 Rag2-deficient mice. After 5-10 weeks, FACS analysis of spleen or peritoneal cavity (PerC) cells of FL (Fig 3A) or BM (Fig 4A) reconstituted mice showed that expression of donor $CD45.1^+$was restricted to $CD19^+$ B cells (first column) and in both cases, all $CD19^+$ cells were IgM^+ (second column). In figure 3, WT cytograms are shown in the first row. Based on CD5 expression, FL-derived pro-B cells propagated with or without stroma gave rise to a large fraction of B1a B cells especially in the PerC (Fig.3A third column cytograms). As expected, BM-derived pro-B cells generated few $CD5^+$ B cells (Fig.4A third column cytograms). Using combined expression of CD21 and CD23 to define $CD21^+/CD23^-$ marginal zone B (MZB) and

$CD21^+/CD23^+$ follicular B cells (FB) spleen $CD19^+$ B cells derived from FL pro-B cells propagated with or without stroma were similar (Fig3A fourth column cytograms). Thus 40 – 60% were FB and 30 – 40% were MZB. The $CD19^+/CD21^-/CD23^-$ cells (lower left quadrant) most likely represent B1 B cells. Immunohistochemical analysis (Fig 3B) of reconstituted mice showed the typical B cell follicular structure seen in WT mice comprising an outer ring of IgM^{high} (green) IgD^{low} (blue) MZB cells surrounding metallophilic macrophages (red) with $IgM^{positive}IgD^{high}$ FB inside. Mice reconstituted with BM pro-B cells cultured by the two methods also showed no obvious differences in CD21 and CD23 expression (Fig4A fourth column cytograms). However, unlike FL-derived pro-B cells, >70% of BM pro-B cell-derived B cells were MZB and only 15 – 20% FB (Fig3B fourth column cytograms).

Spleens of FL-derived pro-B cell reconstituted mice contained around 5×10^6 B cells with no difference between cells cultured with or without stroma (Fig.4B). In contrast, spleens of mice reconstituted with BM-derived pro-B cells contained only 0.5×10^6 B cells irrespective of whether they had been propagated with or without stroma. Thus FL-derived pro-B cells seem to be much more efficient at reconstitution than BM-derived cells.

To test this more stringently, competitive reconstitution experiments were performed. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice were reconstituted with a 1:1 mixture of 5×10^6 CD45.2 FL-derived and CD45.1 BM-derived pro-B cells. After 8 weeks, FACS analysis of one representative mouse (Fig 4C) showed that 44% of splenocytes were $CD19^+$ of which 97.5% (42.9/44) were FL and 2.5% (1.1/4.4) BM-derived. Similar results were obtained in more than three independent experiments. Thus FL-derived pro-B cells are superior to BM-derived ones also in competitive transplantation settings.

In order to identify genes that might be responsible for these observed differences between FL- and BM-derived pro-B cells we performed RNA-sequencing analysis of the two populations. This analysis identified 218 genes differentially expressed more than 2-fold and with high significance (Table SIV). We identified the gene *Lin28B*, as the one most highly expressed in FL-derived pro-B cells compared to BM-derived pro-B (34-fold). It was recently shown that enforced *Lin28B* gene expression in adult HSC converted them functionally into FL HSC [31] and that BM-

derived pro-B cells gained FL-derived pro-B cell properties [32, 33]. To test whether *Lin28B* expression also influenced reconstitution efficiency, we introduced a MigR1 retrovirus encoding *Lin28B* into BM-derived pro-B cells propagated on OP9 stromal cells for two weeks. Then 5 x 10^6 *Lin28B* transduced and 5 x 10^6 non-transduced BM-derived pro-B cells were co-transferred into sub-lethally irradiated Rag2-deficient recipients. Only 2% $CD19^+$ cells were found in the spleens of these mice of which half expressed *Lin28B* (GFP) (Figure S2). Thus *Lin28B* expression did not improve the in vivo B-cell generating capacity of BM-derived pro-B cells. However, practically all $CD19^+$ cells were IgM^+ and > 80% *Lin28B*$^+$ cells were $CD5^+$, thereby resembling B1 B cells. Concerning CD5 expression, *Lin28B* over-expression also converted BM pro-B cells to ones phenotypically resembling FL cells.

Pro-B cell derived B cells mount a T-cell independent immune response

To test whether the B-cell compartments of Rag2-deficient reconstituted mice were functional, 8 weeks after cell transfer they were immunized with the T-cell independent antigen NIP-Ficoll. Serum anti-NIP titers were determined one week before and two weeks after immunization. All reconstituted mice mounted a good IgM anti-NIP response with IgM titers comparable to those in immunized wild type B6 mice (Fig.5A). Mice reconstituted with BM-derived pro-B cells showed a rather low, but significant, IgG anti-NIP response (Fig.5B). However, the IgG anti-NIP response of mice reconstituted with FL pro-B cells was as high or even higher than wild type B6 mice (Fig.5B). Thus the pro-B cell reconstituted B-cell compartments were functional.

The surprisingly high IgG anti-NIP titers in FL pro-B reconstituted mice prompted us to determine their isotypes. Thus, the IgG_{2A} and IgG_{2B} anti-NIP titers were low (Fig 5C) and comparable to those in immunized wild type B6 mice whereas IgG_1 and IgG_3 titers were considerably higher. Although high IgG_3 titers are observed in other T-cell independent responses, switching to IgG_1 was thought to be a highly T-cell dependent phenomenon requiring IL-4 [34-36]. In addition to T cells, mast cells [37, 38], basophils [39], eosinophils [40] and ILC2s [41] may produce IL-4. ILC generation is largely dependent on IL-7 and Rag2/common gamma chain (Rag2cγ) double-deficient mice are practically devoid of ILCs [42, 43]. Therefore, Rag2 and

Rag2cγ double-deficient mice were reconstituted with 10^7 FL-derived pro-B cells and immune responses analyzed as above. Both Rag2 and Rag2cγ double-deficient mice showed a very significant IgM and IgG anti-NIP response (Fig.6A and B), similar and/or even higher than wild type B6 mice. Moreover, both types of reconstituted mice showed a very strong IgG_1 and IgG_3 anti-NIP response (Fig.6 C and D). Thus the IgG class switching observed in immunized reconstituted mice does not seem to be regulated by ILCs.

Reconstitution of the adaptive immune system by *in vitro*-propagated pro-B cells and pro-T cells

Recently, we described a stromal cell free culture system for the long-term propagation of pro-T cells that could be used to reconstitute the T-cell compartment of T-cell deficient mice [22]. However, reconstituted mice contained few regulatory T cells and developed a wasting disease preventable by the co-transfer of mature Treg cells or co-transfer of pro-T cells transduced with a retrovirus encoding *Foxp3-IRES-GFP*. Simultaneous reconstitution with a mixture of *in vitro*-propagated pro-B cells and a non-transduced and *Foxp3* transduced pro-T cells in a 4:1 ratio resulted in T but no B-cell reconstitution (not shown). Therefore, sub-lethally irradiated B6 Rag2-deficient mice were first reconstituted with 10^7 FL-derived pro-B cells and 4 weeks later with 10^7 pro-T cells of which 2.5 x 10^6 were *Foxp3*$^+$. Six weeks after pro-T cell-transfer, a significant B and T-cell reconstitution was seen in peripheral blood cells (data not shown). Because FL pro-B cells partially reconstituted the B-cell compartment of B-cell deficient, T-cell proficient, μMt mice and mounted a T-cell dependent immune response (Figure S3) we tested whether the established B and T cells were functional and could cooperate. Reconstituted mice were therefore immunized with the T-cell dependent antigen NIP-OVA. Rag2-deficient mice reconstituted with pro-B cells alone were used as controls and serum anti-NIP titers determined one week before and two weeks after immunization. No IgM or IgG anti-NIP response was observed in control pro-B cell-reconstituted Rag2-deficient mice (Fig.7A). However, mice reconstituted with both pro-B and pro-T cells mounted a strong IgM and relatively weak, but significant, IgG anti-NIP response (Fig.7A). Thus the reconstituted adaptive immune system was functional.

At 10 – 14 weeks after pro-T cell-transfer, the extent of reconstitution was assessed by FACS analysis. In one such mouse, spleen lymphocytes comprised 10% B cells, 10% CD8 T cells and 20% CD4 T cells (Fig.7B) with ~10% CD4 T cells expressing GFP, indicating they were Tregs. Spleen cell numbers of 7 individual mice 14 weeks after pro-T cell-transfer (Fig.7C) show they contained around 3.5 x 10^6 $CD19^+$ IgM^+ B cells, 1 x 10^6 CD8 T cells and 3 x 10^6 CD4 T cells of which 0.5 x 10^6 were GFP positive; derived from *Foxp3* transduced pro-T cells. Thus transfer of *in vitro*-propagated pro-B and pro-T cells into Rag2-deficient mice resulted into reconstitution of a small, but functionally active, adaptive immune compartment.

DISCUSSION

To our knowledge, this is the first report describing the long-term propagation of BM-derived pro-B cells and their use in reconstituting a functional immune system in immune-deficient recipients. Pro-B cells from $Pax5^{-/-}$ mice were previously shown to have multi-lineage developmental potential [44] and we identified an equivalent $B220^{+}CD117^{+}CD19^{-}NK1.1^{-}$ EPLM cell with lymphoid and myeloid developmental potential (called EPLM) in the BM of WT mice [21]. Herein we show that BM-derived EPLM cultured on OP9 stromal cells plus IL-7 differentiated into $CD19^{+}$, $CD79a^{+}$, $CD79b^{+}$, $Igll1^{+}$ and $Vpreb1^{+}$ pro-B cells with rearranged D_H-J_H genes. Unlike freshly isolated $CD19^{+}$ $CD117^{+}$ pro-B cells, those derived from EPLM could be cultured long-term *in vitro*. Gene expression profiling did not reveal an obvious explanation for this difference in growth capacity but could possibly be regulated via RNA modification or at the translational level [45-47].

To date, long-term growth of pro-B cells required co-culture on stromal cells, yet the specific role of stromal cells in this culture system was unknown. Clearly IL-7, IL-7R (CD127), SCF and its receptor (CD117) as well as FLT3 and its receptor (CD135) all play a role in B-cell development and/or pro-B cell growth [11, 12, 15, 16, 18, 19]. Indeed, we now show that EPLMs efficiently grow and differentiate into pro-B cells when cultured in the combined presence of IL-7, SCF and FLT3L without contact with stromal cells. This indirectly suggests that SCF and FLT3 can substitute for stromal cells.

Pro-B cells maintained on stromal cells plus IL-7 could reconstitute the B-cell compartment of immunodeficient mice [20] and herein we extend these findings. Both FL and BM EPLM-derived pro-B cells propagated either on stromal cells or stromal cell free generated a significant B-cell compartment upon transfer into Rag2-deficient mice. However, FL-derived pro-B cells were about 40 fold more efficient at B-cell reconstitution than their BM-derived partners.

Although pro-B cells expressing the *Lin28B* can acquire characteristics of FL-derived pro-B cells [32] and *in vitro*-propagated *Lin28B*-expressing BM-derived pro-B cells could generate CD5 positive B cells in vivo, there was no improvement in the in vivo-reconstitution capacity of BM pro-B cells.

An enlarged MZB compartment is frequently observed in B lymphopenic mice [48-52] suggesting that newly-formed B cells first fill the MZB compartment. This phenomenon could be due either to a specialized environment within the splenic marginal zone or that only B cells expressing certain BCR home to and expand in this anatomical location [53].

When immunized with the T-independent antigen NIP-Ficoll, all reconstituted mice mounted an anti-NIP response indicating their B cell compartments were functional.. Mice reconstituted with BM-derived pro-B cells had a slightly lower IgM anti-NIP titer than WT mice possibly due to poor B cell reconstitution and showed a significant IgG anti-NIP response although still at least 10 fold lower than WT mice. In marked contrast, Rag2-deficient mice reconstituted with FL pro-B cells mounted an IgM and IgG anti-NIP response that was identical, or even higher than, WT mice suggesting that fetal progenitor-derived B cells are the main anti-NIP responders in a primary immunization. In B6 mice, the primary anti-NIP antibodies use predominantly λ1 light chains [54-56]. The anti-NIP response of pro-B cell-reconstituted mice was likewise dominated by λ1-containing antibodies (data not shown) indicating that the anti-NIP B-cell repertoires in pro-B cell reconstituted mice are identical to those of WT mice.

Surprisingly, FL pro-B cell-reconstituted mice had a high IgG_1 anti-NIP response an antibody class normally associated with IL-4. Since reconstituted mice were devoid of T cells, the question arose which cell type was responsible for IL-4 production [34-36]. Pro-B cell reconstituted Rag2cγ double-deficient mice, deficient in ILCs, also mounted high IgG_1 response suggesting that ILCs were not the source of IL-4 [42, 43]. Mast cells, basophils and/or eosinophils, could be the source of IL-4 [37-40]. FL pro-B cells were also able to reconstitute a functional B-cell compartment in T-cell containing, B-cell deficient μMt mice (Figure S3). This result might seem surprising given that the T-cell compartments in these mice may not be tolerant to mature B cells.

Pro-T cells propagated *in vitro* under stromal free conditions can reconstitute the T-cell compartments of T-cell deficient mice [22]. Here we show that the combined transfer of stromal cell free-propagated pro-B and pro-T cells into Rag2-deficient mice results in the generation of a functional adaptive immune repertoire

capable of mounting a T-dependent antibody response. For efficient reconstitution, pro-T cell transfer had to be performed 2-3 weeks after the pro-B cell reconstitution suggesting that when transferred together, T cells develop that are not tolerant to B cells and therefore eliminate them.

Overall the findings described herein show that progenitor lymphocytes can be readily propagated under stromal free conditions *in vitro* and that these cells can be used to reconstitute mice with mature functional lymphocytes. Based on these results the establishment of stromal cell free culture systems for human lymphocyte progenitors might be of great interest since unlike stromal cell-propagated lymphocyte progenitors, stromal cell free propagated cells could be potentially used for therapeutic purposes in patients with B and/or T-cell deficiencies.

MATERIALS AND METHODS

Mice

Female C57BL/6 CD45.1 and CD45.2, C57BL/6 Rag2-deficient [23], and C57BL/6 Bcl-2 transgenic mice [24] with 5–8 weeks of age were used. The appearance of vaginal plugs was counted as day 0 of gestation and embryos were taken at day E17.5. All mice were bred and maintained in our animal facility under specific pathogen-free conditions. Animal experiments were carried out within institutional guidelines (authorization numbers 1886 and 1888 from Kantonales Veterinäramt, Basel).

Cell lines, cell culture, and supplements

The OP9 stromal cell line [25] was cultured as a monolayer in IMDM supplemented with 2% FBS, 5×10^{-5} M β-mercaptoethanol, 1 mM glutamine, 0.03% w/v Primatone (Quest, Naarden, The Netherlands), and 100 U/mL penicillin. For pro-B cell culture, $CD117^{+}B220^{+}CD19^{-}NK1.1^{-}$ cells were sorted from the FL or the femoral BM of adult mice and cultured at 10^{4}/ml in supplemented IMDM either on a semi-confluent layer of 30 Gy γ-irradiated OP9 stromal cells in the presence of 100U/ml IL-7, or without stromal cell support but in the presence of 100U/ml IL-7, 50ng/ml FLT3L, and 100ng/ml SCF. IL-7 was derived from culture supernatant of J558L cells transfected with murine IL-7 cDNA. Polyhistidin-tagged SCF was purified from transfected Rosetta pLacI bacteria using Ni-NTA-agarose beads (Qiagen, Venlo, NL). A vector expressing a human FLT3L-Fc fusion protein was expressed in Chinese hamster ovary cells. The supernatant was passed over a protein A-Sepharose (GE Healthcare, Chalfont St. Giles, GB) column in order to purify the protein. Pro T cells were cultured as previously described [22].

Antibodies, flow cytometry, and sorting

FITC-, PE-, allophycocyanin-, or biotin-labeled mAbs specific for CD117, B220, CD19, NK1.1, IgM, Igκ, CD5, CD45.1, CD21, CD23, CD4, CD8α, and TCRβ were either purchased from BD Biosciences (Franklin Lakes, NJ, USA) or eBiosciences (San Diego, CA, USA), or purified from hybridoma culture supernatants according to standard procedures. Staining of the cells was performed as described before [26]. Flow

cytometry was done using a FACS Calibur (BD Biosciences) and data were analyzed using the CellQuest Pro (BD Biosciences) or FlowJo Software (Treestar). For cell sorting, a FACSAria IIu (BD Biosciences) was used (>98% purity).

Transfer of cultured progenitor cells

Recipient mice were γ-irradiated using a Cobalt source (Gammacell 40, Atomic Energy of Canada, Ltd) 4h prior to reconstitution. The indicated number of pro-B or pro T cells was then injected into the tail vein.

Immunohistochemical analysis

To analyze pro-B-cell derived B-cell localization in the spleen, the 5µm snap frozen and acetone-fixed sections were incubated with FITC-labelled anti-IgM (clone M41, self-made), APC-labelled anti-IgD (clone 1.19 self-made), and biotinylated anti-MOMA-1 (Vector, Burlingame, CA), which was revealed with PE-conjugated streptavidin (SouthernBiotech, Birmingham, AL). Confocal microscopy images were taken with a LSM 510 Meta (Zeiss, Oberkochen, D) and analyzed using the ImageJ software and the Fiji image processing package.

NIP-specific antibody responses

Reconstituted mice were immunized subcutaneously with 100µg NIP-Ficoll or NIP-OVA in a 1:1 CFA emulsion. Serum IgM and IgG against NIP was analyzed at day 14 using ELISA as described in [27].

PCR analyses

The PCR conditions for amplifying D_HJ_H rearrangements were described elsewhere [4, 6, 28]. The primers used were D_H 5′-TTCAAAGCACAATGCCTGGCT-3′ and J_H3 5′-GTCTAGATTCTCACAAGAGTCCGATAGACCCTGG-3′. Oligonucleotide primers used for CD79A verification from pro-B cDNA were 5′-TGTTTGGGTCCCGGATGCCA-3′and 5′-CACGCGGAGGTAAGTACCACA-3′, for CD79B 5′-TCTTCTCAGGTGAGCCGGTA-3′and 5′-TATGGTTGGCGCTGTCACAT-3′, for IGLL1 5′- AGTAGGACAGACTCTGGGCA-3′and 5′-GGCTGACCTAGGATTGTGAGC-3′, for VPREB1 5′-CTCCGGGTCCAAAGATACGAC-3′and 5′-GCTCATAGCAACACCGCAGAA-3′, and for beta-actin

5´-GAAGTCTAGAGCAACATAGCACAGCTTCTC-3´and

5´-GTGGGAATTCGTCAGAAGGACTCCTATGTG-3´.

Statistical analysis

Statistical analysis was performed with Prism 6.0g software (GraphPad Software, Inc.). Two-tailed unpaired Student *t* tests were used for statistical comparisons. If not differently indicated, data are presented as mean values ± SEM from three independent experiments. ns not significant or $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, ****$P \leq 0.0001$.

REFERENCES

1 **Rolink, A., Grawunder, U., Winkler, T. H., Karasuyama, H. and Melchers, F.,** IL-2 receptor alpha chain (CD25, TAC) expression defines a crucial stage in pre-B cell development. *Int Immunol* 1994. **6**: 1257-1264.

2 **Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D. and Hayakawa, K.,** Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J Exp Med* 1991. **173**: 1213-1225.

3 **Rolink, A. and Melchers, F.,** Molecular and cellular origins of B lymphocyte diversity. *Cell* 1991. **66**: 1081-1094.

4 **ten Boekel, E., Melchers, F. and Rolink, A.,** The status of Ig loci rearrangements in single cells from different stages of B cell development. *Int Immunol* 1995. **7**: 1013-1019.

5 **Nutt, S. L., Heavey, B., Rolink, A. G. and Busslinger, M.,** Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 1999. **401**: 556-562.

6 **Nutt, S. L., Thevenin, C. and Busslinger, M.,** Essential functions of Pax-5 (BSAP) in pro-B cell development. *Immunobiology* 1997. **198**: 227-235.

7 **Morrison, A. M., Nutt, S. L., Thevenin, C., Rolink, A. and Busslinger, M.,** Loss- and gain-of-function mutations reveal an important role of BSAP (Pax-5) at the start and end of B cell differentiation. *Semin Immunol* 1998. **10**: 133-142.

8 **Kozmik, Z., Wang, S., Dorfler, P., Adams, B. and Busslinger, M.,** The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. *Mol Cell Biol* 1992. **12**: 2662-2672.

9 **Rolink, A., Haasner, D., Nishikawa, S. and Melchers, F.,** Changes in frequencies of clonable pre B cells during life in different lymphoid organs of mice. *Blood* 1993. **81**: 2290-2300.

10 **Urbanek, P., Wang, Z. Q., Fetka, I., Wagner, E. F. and Busslinger, M.,** Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* 1994. **79**: 901-912.

11 **Peschon, J. J., Morrissey, P. J., Grabstein, K. H., Ramsdell, F. J., Maraskovsky, E., Gliniak, B. C., Park, L. S., Ziegler, S. F., Williams, D. E., Ware, C. B., Meyer, J. D. and Davison, B. L.,** Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J Exp Med* 1994. **180**: 1955-1960.

12 **von Freeden-Jeffry, U., Vieira, P., Lucian, L. A., McNeil, T., Burdach, S. E. and Murray, R.,** Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J Exp Med* 1995. **181**: 1519-1526.

13 **Agosti, V., Corbacioglu, S., Ehlers, I., Waskow, C., Sommer, G., Berrozpe, G., Kissel, H., Tucker, C. M., Manova, K., Moore, M. A., Rodewald, H. R. and Besmer, P.,** Critical role for Kit-mediated Src kinase but not PI 3-kinase signaling in pro T and pro B cell development. *J Exp Med* 2004. **199**: 867-878.

14 **Billips, L. G., Petitte, D., Dorshkind, K., Narayanan, R., Chiu, C. P. and Landreth, K. S.,** Differential roles of stromal cells, interleukin-7, and kit-ligand in the regulation of B lymphopoiesis. *Blood* 1992. **79**: 1185-1192.

15 **Mackarehtschian, K., Hardin, J. D., Moore, K. A., Boast, S., Goff, S. P. and Lemischka, I. R.,** Targeted disruption of the flk2/flt3 gene leads to

deficiencies in primitive hematopoietic progenitors. *Immunity* 1995. **3**: 147-161.

16 **McKenna, H. J., Stocking, K. L., Miller, R. E., Brasel, K., De Smedt, T., Maraskovsky, E., Maliszewski, C. R., Lynch, D. H., Smith, J., Pulendran, B., Roux, E. R., Teepe, M., Lyman, S. D. and Peschon, J. J.,** Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* 2000. **95**: 3489-3497.

17 **Tsapogas, P., Swee, L. K., Nusser, A., Nuber, N., Kreuzaler, M., Capoferri, G., Rolink, H., Ceredig, R. and Rolink, A.,** In vivo evidence for an instructive role of fms-like tyrosine kinase-3 (FLT3) ligand in hematopoietic development. *Haematologica* 2014. **99**: 638-646.

18 **Rolink, A., Streb, M., Nishikawa, S. and Melchers, F.,** The c-kit-encoded tyrosine kinase regulates the proliferation of early pre-B cells. *Eur J Immunol* 1991. **21**: 2609-2612.

19 **Buza-Vidas, N., Cheng, M., Duarte, S., Nozad, H., Jacobsen, S. E. and Sitnicka, E.,** Crucial role of FLT3 ligand in immune reconstitution after bone marrow transplantation and high-dose chemotherapy. *Blood* 2007. **110**: 424-432.

20 **Rolink, A., Kudo, A., Karasuyama, H., Kikuchi, Y. and Melchers, F.,** Long-term proliferating early pre B cell lines and clones with the potential to develop to surface Ig-positive, mitogen reactive B cells in vitro and in vivo. *EMBO J* 1991. **10**: 327-336.

21 **Balciunaite, G., Ceredig, R., Massa, S. and Rolink, A. G.,** A B220+ CD117+ CD19- hematopoietic progenitor with potent lymphoid and myeloid developmental potential. *Eur J Immunol* 2005. **35**: 2019-2030.

22 **Gehre, N., Nusser, A., von Muenchow, L., Tussiwand, R., Engdahl, C., Capoferri, G., Bosco, N., Ceredig, R. and Rolink, A. G.,** A stromal cell free culture system generates mouse pro-T cells that can reconstitute T-cell compartments in vivo. *Eur J Immunol* 2015. **45**: 932-942.

23 **Shinkai, Y., Rathbun, G., Lam, K. P., Oltz, E. M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A. M. and et al.,** RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 1992. **68**: 855-867.

24 **Domen, J., Gandy, K. L. and Weissman, I. L.,** Systemic overexpression of BCL-2 in the hematopoietic system protects transgenic mice from the consequences of lethal irradiation. *Blood* 1998. **91**: 2272-2282.

25 **Nakano, T., Kodama, H. and Honjo, T.,** Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* 1994. **265**: 1098-1101.

26 **Rolink, A., ten Boekel, E., Melchers, F., Fearon, D. T., Krop, I. and Andersson, J.,** A subpopulation of B220+ cells in murine bone marrow does not express CD19 and contains natural killer cell progenitors. *J Exp Med* 1996. **183**: 187-194.

27 **Schubart, K., Massa, S., Schubart, D., Corcoran, L. M., Rolink, A. G. and Matthias, P.,** B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat Immunol* 2001. **2**: 69-74.

28 **Gu, H., Kitamura, D. and Rajewsky, K.,** B cell development regulated by gene rearrangement: arrest of maturation by membrane-bound D mu protein and selection of DH element reading frames. *Cell* 1991. **65**: 47-54.

29 **Reininger, L., Radaszkiewicz, T., Kosco, M., Melchers, F. and Rolink, A. G.,** Development of autoimmune disease in SCID mice populated with long-term "in vitro" proliferating (NZB x NZW)F1 pre-B cells. *J Exp Med* 1992. **176**: 1343-1353.

30 **Reininger, L., Winkler, T. H., Kalberer, C. P., Jourdan, M., Melchers, F. and Rolink, A. G.,** Intrinsic B cell defects in NZB and NZW mice contribute to systemic lupus erythematosus in (NZB x NZW)F1 mice. *J Exp Med* 1996. **184**: 853-861.

31 **Yuan, J., Nguyen, C. K., Liu, X., Kanellopoulou, C. and Muljo, S. A.,** Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. *Science* 2012. **335**: 1195-1200.

32 **Zhou, Y., Li, Y. S., Bandi, S. R., Tang, L., Shinton, S. A., Hayakawa, K. and Hardy, R. R.,** Lin28b promotes fetal B lymphopoiesis through the transcription factor Arid3a. *J Exp Med* 2015. **212**: 569-580.

33 **Li, Y. S., Zhou, Y., Tang, L., Shinton, S. A., Hayakawa, K. and Hardy, R. R.,** A developmental switch between fetal and adult B lymphopoiesis. *Ann N Y Acad Sci* 2015. **1362**: 8-15.

34 **Moon, H. B., Severinson, E., Heusser, C., Johansson, S. G., Moller, G. and Persson, U.,** Regulation of IgG1 and IgE synthesis by interleukin 4 in mouse B cells. *Scand J Immunol* 1989. **30**: 355-361.

35 **Snapper, C. M., Finkelman, F. D. and Paul, W. E.,** Differential regulation of IgG1 and IgE synthesis by interleukin 4. *J Exp Med* 1988. **167**: 183-196.

36 **Snapper, C. M. and Paul, W. E.,** Interferon-gamma and B cell stimulatory factor-1 reciprocally regulate Ig isotype production. *Science* 1987. **236**: 944-947.

37 **Plaut, M., Pierce, J. H., Watson, C. J., Hanley-Hyde, J., Nordan, R. P. and Paul, W. E.,** Mast cell lines produce lymphokines in response to cross-linkage of Fc epsilon RI or to calcium ionophores. *Nature* 1989. **339**: 64-67.

38 **Bradding, P., Feather, I. H., Howarth, P. H., Mueller, R., Roberts, J. A., Britten, K., Bews, J. P., Hunt, T. C., Okayama, Y., Heusser, C. H. and et al.,** Interleukin 4 is localized to and released by human mast cells. *J Exp Med* 1992. **176**: 1381-1386.

39 **Min, B., Prout, M., Hu-Li, J., Zhu, J., Jankovic, D., Morgan, E. S., Urban, J. F., Jr., Dvorak, A. M., Finkelman, F. D., LeGros, G. and Paul, W. E.,** Basophils produce IL-4 and accumulate in tissues after infection with a Th2-inducing parasite. *J Exp Med* 2004. **200**: 507-517.

40 **Nonaka, M., Nonaka, R., Woolley, K., Adelroth, E., Miura, K., Okhawara, Y., Glibetic, M., Nakano, K., O'Byrne, P., Dolovich, J. and et al.,** Distinct immunohistochemical localization of IL-4 in human inflamed airway tissues. IL-4 is localized to eosinophils in vivo and is released by peripheral blood eosinophils. *J Immunol* 1995. **155**: 3234-3244.

41 **Doherty, T. A., Khorram, N., Lund, S., Mehta, A. K., Croft, M. and Broide, D. H.,** Lung type 2 innate lymphoid cells express cysteinyl leukotriene receptor 1,

which regulates TH2 cytokine production. *J Allergy Clin Immunol* 2013. **132**: 205-213.

42 **Chappaz, S. and Finke, D.,** The IL-7 signaling pathway regulates lymph node development independent of peripheral lymphocytes. *J Immunol* 2010. **184**: 3562-3569.

43 **Spits, H., Artis, D., Colonna, M., Diefenbach, A., Di Santo, J. P., Eberl, G., Koyasu, S., Locksley, R. M., McKenzie, A. N., Mebius, R. E., Powrie, F. and Vivier, E.,** Innate lymphoid cells--a proposal for uniform nomenclature. *Nat Rev Immunol* 2013. **13**: 145-149.

44 **Rolink, A. G., Nutt, S. L., Melchers, F. and Busslinger, M.,** Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* 1999. **401**: 603-606.

45 **Frye, M., Jaffrey, S. R., Pan, T., Rechavi, G. and Suzuki, T.,** RNA modifications: what have we learned and where are we headed? *Nat Rev Genet* 2016.

46 **Preston, G. C., Sinclair, L. V., Kaskar, A., Hukelmann, J. L., Navarro, M. N., Ferrero, I., MacDonald, H. R., Cowling, V. H. and Cantrell, D. A.,** Single cell tuning of Myc expression by antigen receptor signal strength and interleukin-2 in T lymphocytes. *EMBO J* 2015. **34**: 2008-2024.

47 **Hukelmann, J. L., Anderson, K. E., Sinclair, L. V., Grzes, K. M., Murillo, A. B., Hawkins, P. T., Stephens, L. R., Lamond, A. I. and Cantrell, D. A.,** The cytotoxic T cell proteome and its shaping by the kinase mTOR. *Nat Immunol* 2016. **17**: 104-112.

48 **Martin, F. and Kearney, J. F.,** Positive selection from newly formed to marginal zone B cells depends on the rate of clonal production, CD19, and btk. *Immunity* 2000. **12**: 39-49.

49 **von Muenchow, L., Engdahl, C., Karjalainen, K. and Rolink, A. G.,** The selection of mature B cells is critically dependent on the expression level of the co-receptor CD19. *Immunol Lett* 2014. **160**: 113-119.

50 **Carvalho, T. L., Mota-Santos, T., Cumano, A., Demengeot, J. and Vieira, P.,** Arrested B lymphopoiesis and persistence of activated B cells in adult interleukin 7(-/)- mice. *J Exp Med* 2001. **194**: 1141-1150.

51 **Hao, Z. and Rajewsky, K.,** Homeostasis of peripheral B cells in the absence of B cell influx from the bone marrow. *J Exp Med* 2001. **194**: 1151-1164.

52 **Martin, F. and Kearney, J. F.,** Marginal-zone B cells. *Nat Rev Immunol* 2002. **2**: 323-335.

53 **Chen, X., Martin, F., Forbush, K. A., Perlmutter, R. M. and Kearney, J. F.,** Evidence for selection of a population of multi-reactive B cells into the splenic marginal zone. *Int Immunol* 1997. **9**: 27-41.

54 **Smith, F. I., Tesch, H. and Rajewsky, K.,** Heterogeneous and monoclonal helper T cells induce similar anti-(4-hydroxy-3-nitrophenyl)acetyl (NP) antibody populations in the primary adoptive response. II. Lambda light chain dominance and idiotope expression. *Eur J Immunol* 1984. **14**: 195-200.

55 **Jacob, J., Kassir, R. and Kelsoe, G.,** In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. I. The architecture and dynamics of responding cell populations. *J Exp Med* 1991. **173**: 1165-1175.

56 **Jack, R. S., Imanishi-Kari, T. and Rajewsky, K.,** Idiotypic analysis of the response of C57BL/6 mice to the (4-hydroxy-3-nitrophenyl)acetyl group. *Eur J Immunol* 1977. **7**: 559-565.

FIGURES

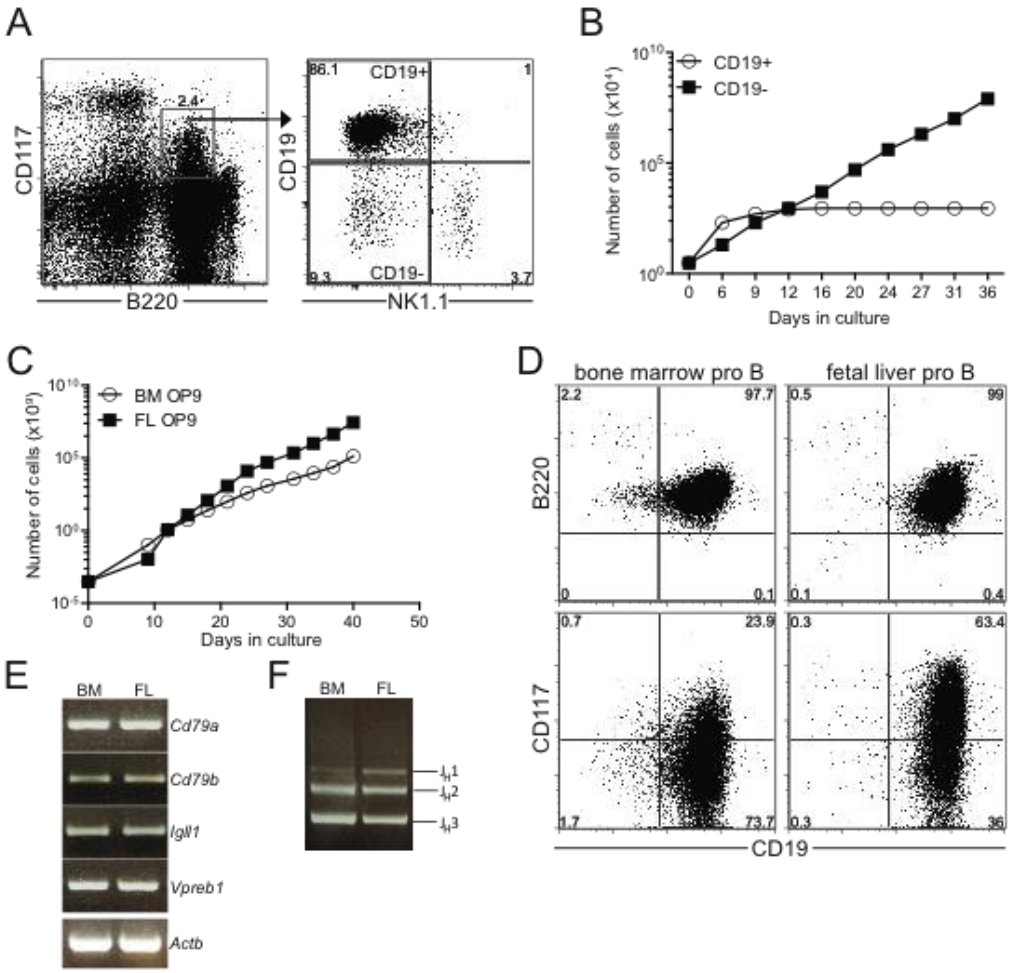


Figure 1. Establishment of EPLM-derived long-term proliferating pro-B cell lines from BM and FL. FL and BM-derived $CD19^{-}B220^{+}CD117^{+}NK1.1^{-}$ cells were sorted and maintained on OP9 stromal cells plus IL-7 (A) Gating strategy for sorting of $CD19^{+}$ and $CD19^{-}$ $B220^{+}CD117^{+}$ BM cells. (B) In vitro growth capacity of sorted CD19+ and CD19- BM cells on OP9 stromal cells plus IL-7. (C) Comparison of the in vitro growth of FL and BM-derived pro-B cells for about 40 days. (D) Representative FACS plots showing B220, CD19 and CD117 expression in FL and BM-derived pro-B cells cultured for 12 days. (E) RT-PCR analysis of *Cd79a*, *Cd79b*, *Igll1*, *Vpreb1*, and *Actb* expression in FL and BM-derived pro-B cells cultured for 12 days. (F) Genomic D_H-J_H rearrangement analysis of FL and BM-derived pro-B cells cultures for 12 days.

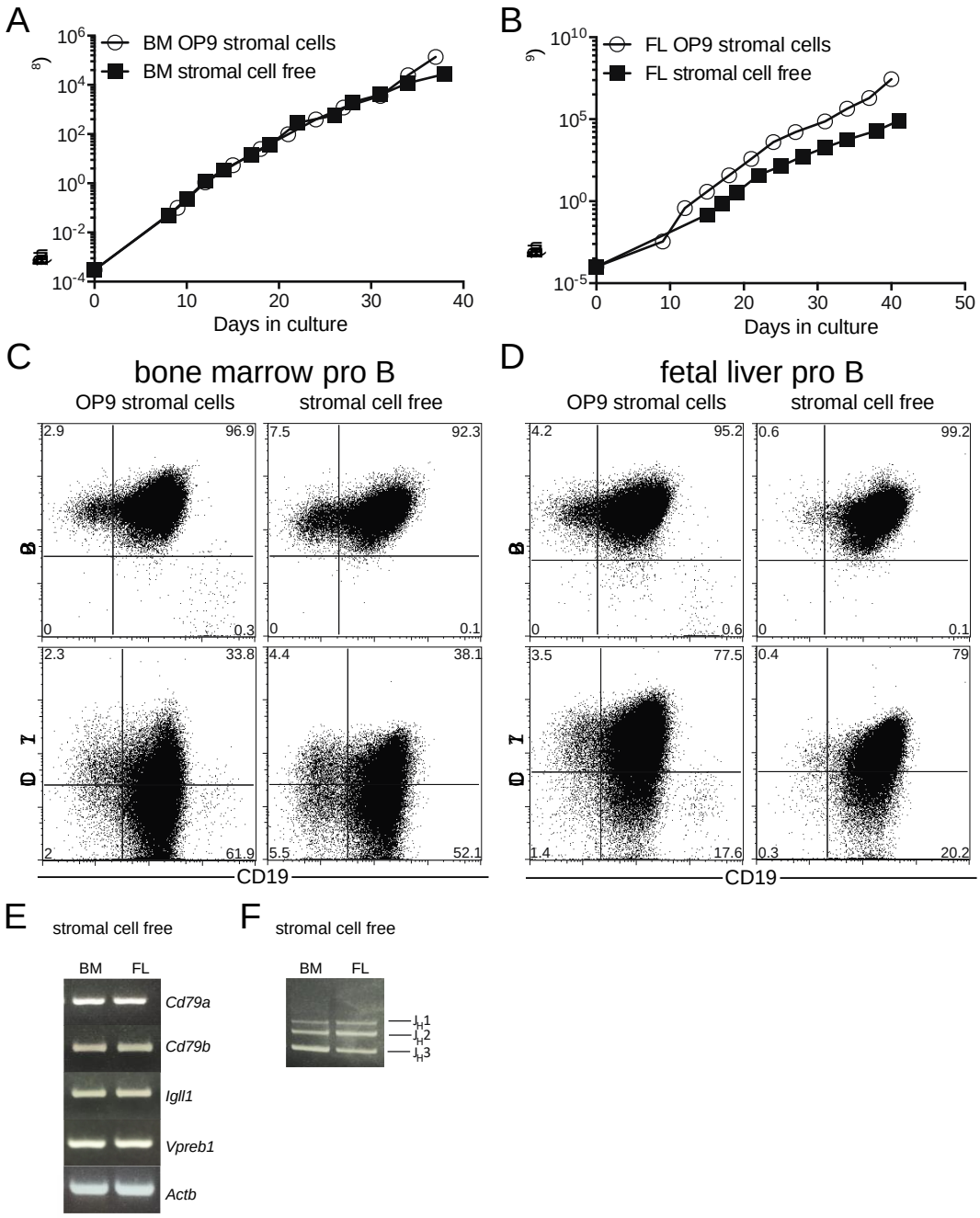


Figure 2. Establishment of EPLM-derived long-term proliferating pro-B cell lines from BM and FL under stromal cell free conditions. BM or FL-derived CD19$^-$ B220$^+$CD117$^+$NK1.1$^-$ cells were sorted and maintained either on OP9 stromal cells plus IL-7 or in the presence of IL-7, SCF and FLT3L without stromal cells. (A) Comparison of the growth capacity of BM-derived pro-B cells cultured in the absence (black squares) or presence (white circles) of stromal cells for about 40 days. (B) Comparison of the growth capacity of FL-derived pro-B cells cultured in the absence (black squares) or presence (white circles) of stromal cells for about 40 days. (C) Representative FACS plots showing B220, CD19 and CD117 expression on BM-

derived pro-B cells cultured for 14 days with or without stromal cells, as indicated. (D) Representative FACS plots showing B220, CD19 and CD117 expression on FL-derived pro-B cells cultured for 14 days with or without stromal cells, as indicated. (E) RT-PCR analysis of *Cd79a*, *Cd78b*, *Igll1*, *Vpreb1* and *Actb* expression in FL and BM-derived pro-B cells cultured for 14 days in the presence of IL-7, SCF and FLT3L without stromal cells. (F) Genomic D_H-J_H rearrangement analysis of FL and BM-derived pro-B cells cultured for 14 days in the presence of IL-7, SCF and FLT3L without stromal cells.

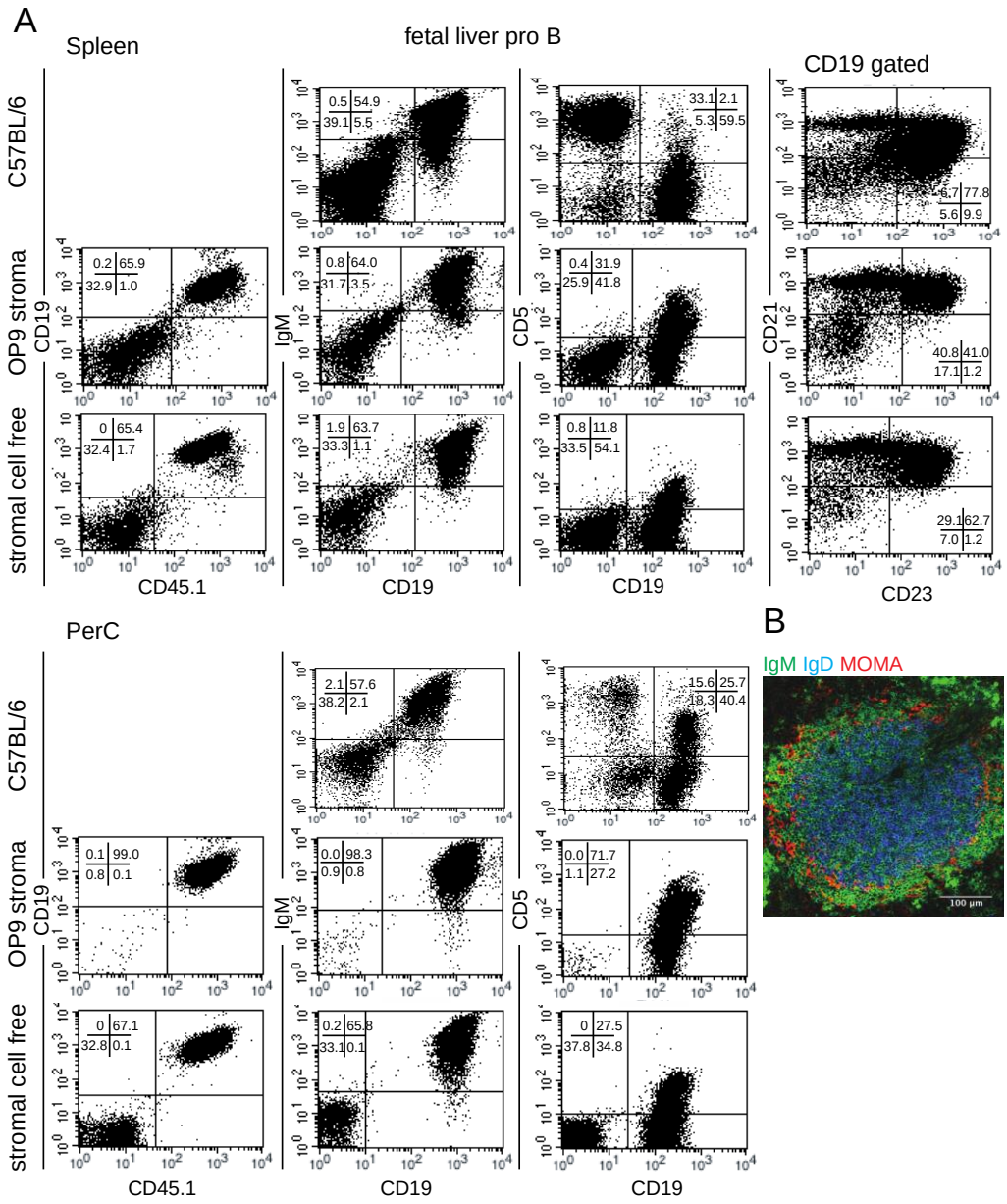


Figure 3. In vivo reconstitution potential of FL-derived pro-B cells propagated in vitro. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice were injected intravenously with 10^7 CD45.1 FL-derived pro-B cells propagated in the presence or absence of stromal cells. (A) Representative FACS analysis of B cell populations in the spleen and peritoneal cavity (PerC) of WT controls and recipient mice 8 weeks after cell transfer of the indicated pro-B cells. (B) Representative picture of a staining for IgM (green), IgD (blue), and metallophilic macrophages (MOMA-1 in red) on a spleen section of mice reconstituted with FL-derived pro-B cells propagated on OP9 stromal cells (20x magnification). In total 28 individual mice were injected: 14 mice with FL-derived pro-B cells propagated with OP9 stromal cells and 14 mice with FL-derived pro-B cells propagated without stromal cells.

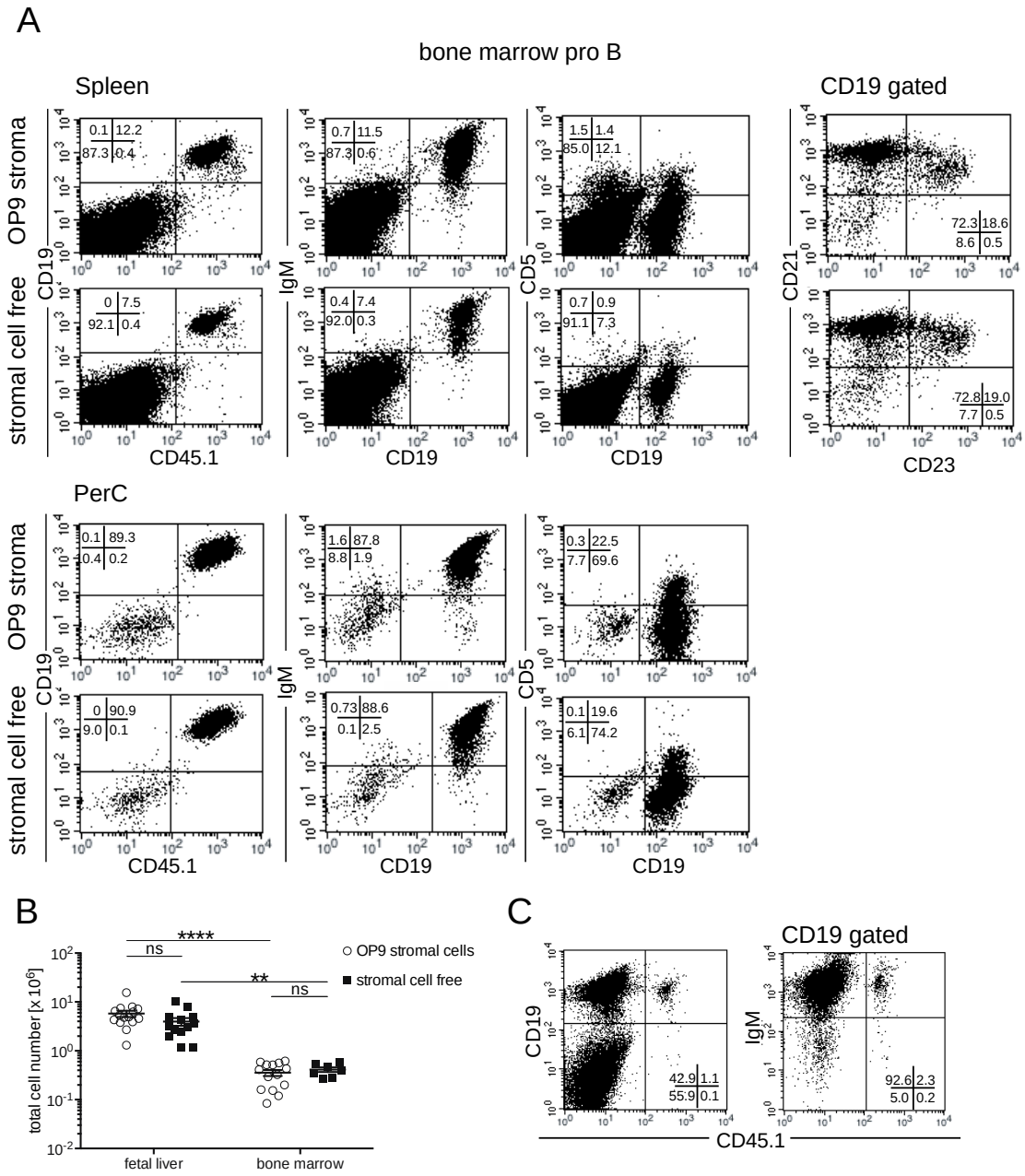


Figure 4. In vivo reconstitution potential of BM-derived pro-B cells propagated in vitro and comparison to the corresponding potential of FL-derived pro-B cells. Sub-lethally irradiated CD45.2 B6 Rag2-deficient mice (n=21) were injected intravenously with 10^7 CD45.1 BM-derived pro-B cells propagated in the presence (14 mice injected) or absence (7 mice injected) of stromal cells. (A) Representative FACS analysis of B-cell populations in the spleen and peritoneal cavity (PerC) of WT controls and recipient mice 8 weeks after cell transfer of the indicated pro-B cells. (B) Absolute numbers of $CD19^+$ IgM^+ cells in the spleens of mice injected with BM-derived pro-B cells and FL-derived pro-B cells (Figure 3). Unpaired student's t-test. Data shown above are mean ± SEM and each circle or square is an individual mouse.

ns: not significant or P > 0.05, **P ≤ 0.01 and ****P ≤ 0.0001. (C) CD19, IgM, and CD45.1 expression from a spleen of a CD45.2 B6 Rag2-deficient mouse 8 weeks after transfer of 5 x 10^6 CD45.2 FL-derived pro-B cells and 5 x 10^6 CD45.1 BM-derived pro-B cells. Both pro-B cells were propagated stromal cell free. A representative FACS analysis out of 3 independently performed experiments is shown.

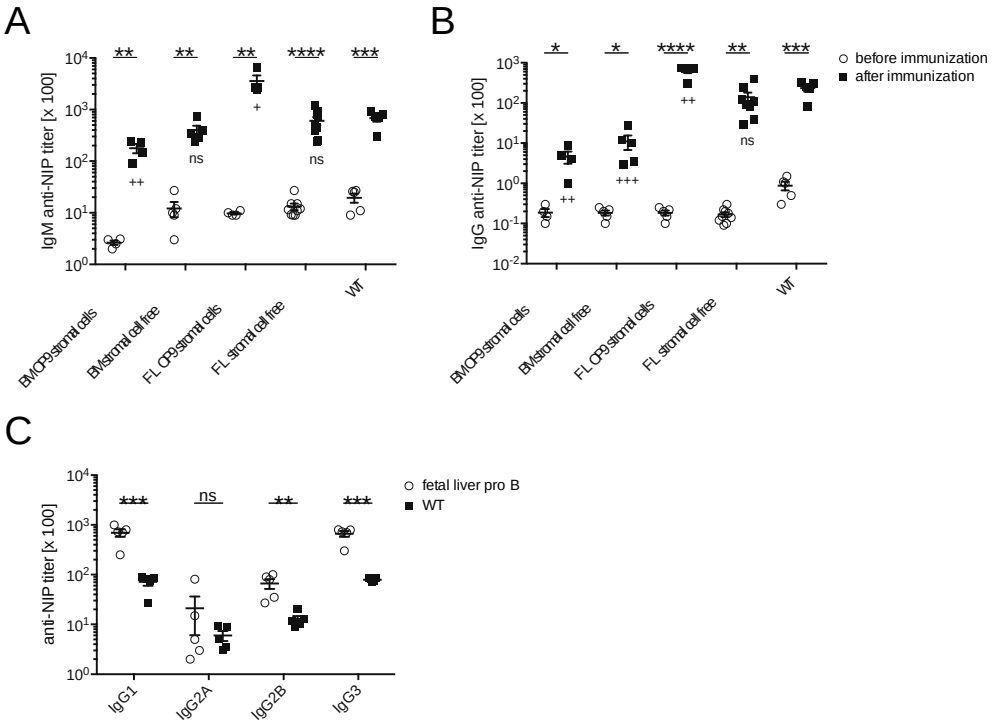


Figure 5. T-cell independent responses of mice reconstituted with pro-B cells propagated in the absence or presence of stromal cells. B6 Rag2-deficient mice were reconstituted with BM or FL derived pro-B cells propagated with or without stromal cells and immunized with NIP-Ficoll 8 weeks after cell transfer. Normal B6 (WT) mice were used as controls and statistical significance compared to WT is indicated below the after immunization data points (black squares). Titers were defined as the serum dilutions that gave 2 times background OD values in the ELISA test. Titers of the indicated antibodies were determined in sera taken 1 week before immunization (A and B open symbols) and 2 weeks after immunization (A and B closed symbols). (A) IgM anti-NIP titers. (B) IgG anti-NIP titers. (C) Anti-NIP IgG subclass titers in the serum of mice reconstituted with FL-derived pro-B cells and WT mice. ns: not significant or $P > 0.05$, * or $^{+}$ $P \leq 0.05$, ** or $^{++}$ $P \leq 0.01$, *** or $^{+++}$ $P \leq 0.001$, **** $P \leq 0.0001$. * represents significance compared to pre-bleed. $^{+}$ represents significance compared to titer of WT mice. Each symbol represents an individual mouse. Unpaired student's t test; n = 5-8. Data shown above are mean ± SEM and two independent experiments were performed.

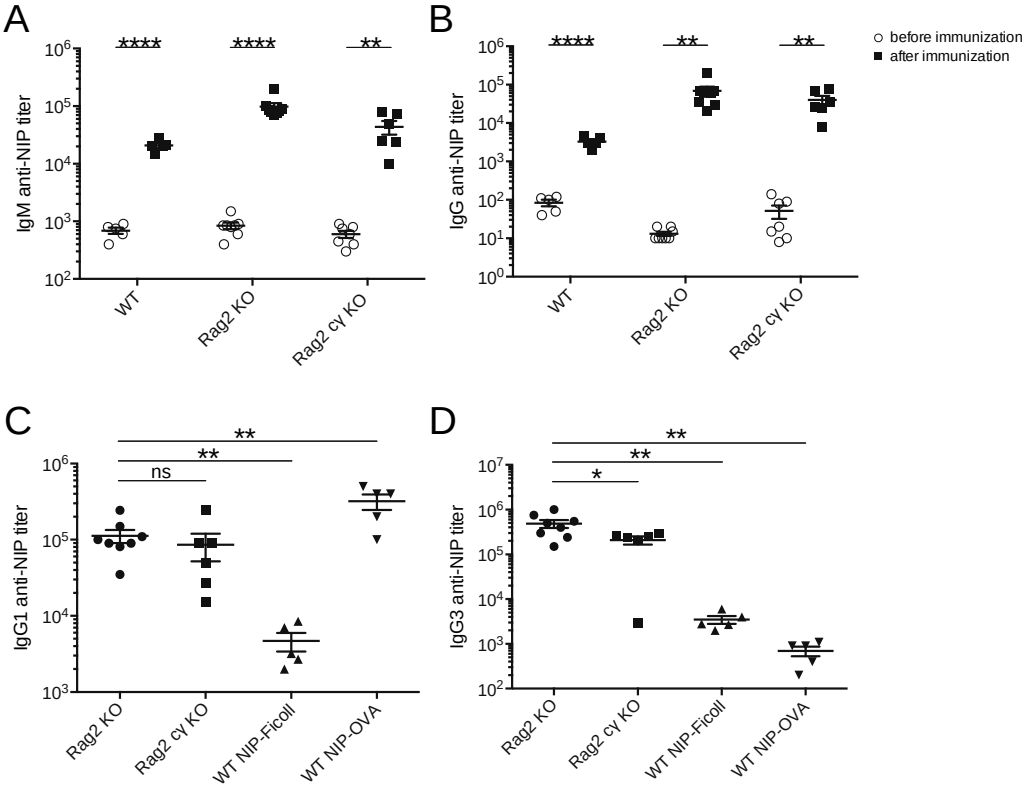


Figure 6. Anti-NIP responses of WT mice, Rag2-deficient and Rag2cγ double-deficient mice reconstituted with FL pro-B cells propagated on stromal cells and immunized with NIP-Ficoll 8 weeks after cell transfer. Titers were defined as the serum dilutions that gave 2 times background OD values in the ELISA test. Titers were determined in sera taken 1 week before immunization (A and B open symbols) and 2 weeks after immunization (A and B closed symbols) (A) IgM anti-NIP titers. (B) IgG anti-NIP titers. (C) IgG1 anti-NIP titers. Sera from B6 mice immunized with NIP-Ficoll or NIP-OVA were used as positive controls. (D) IgG3 anti-NIP titers. Sera from B6 mice immunized with NIP-Ficoll or NIP-OVA were used as positive controls. Collective data from two independent experiments are shown and each symbol represents an individual mouse. ns: not significant, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Unpaired student's t test; n = 5-8. Data shown above are mean ± SEM.

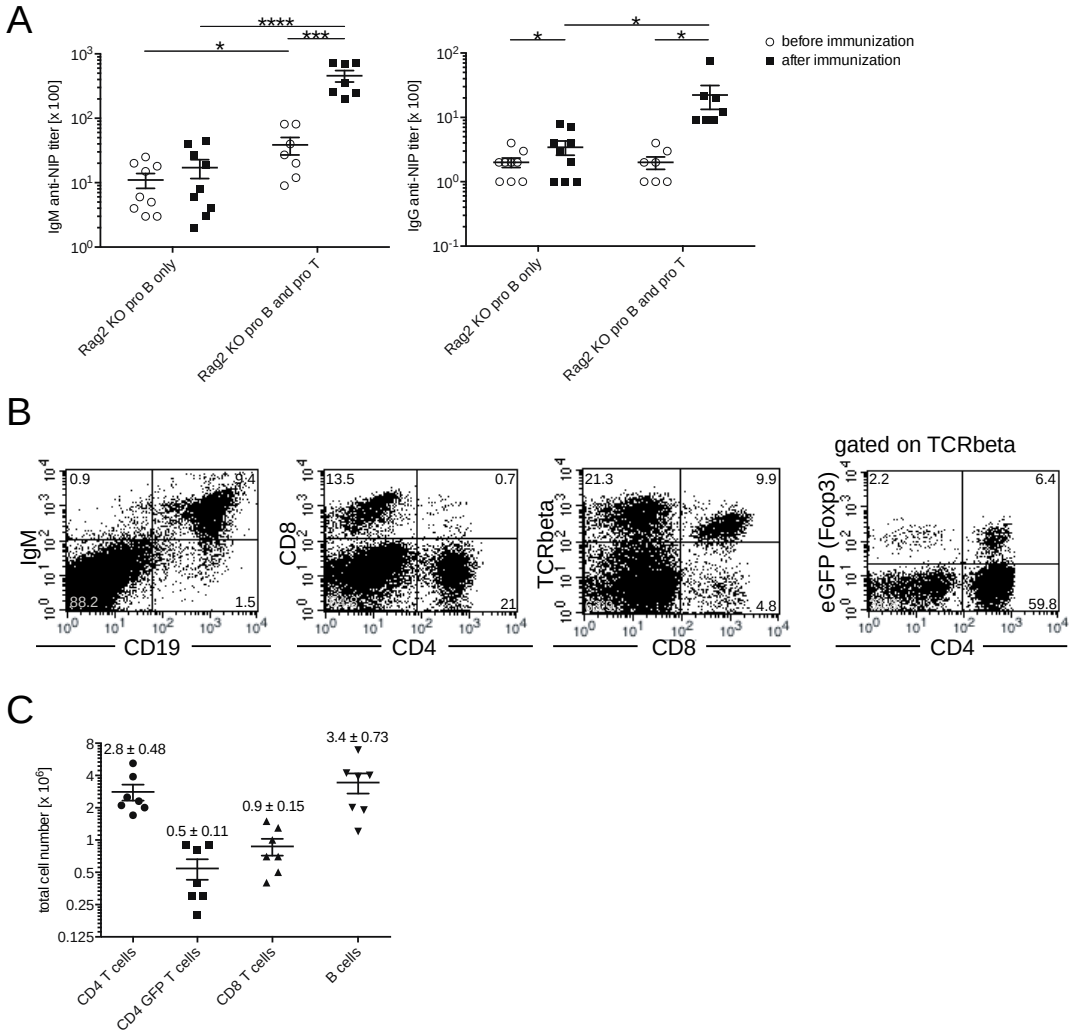


Figure 7. In vivo lymphoid reconstitution and immune responses in mice injected with in vitro propagated pro-B and pro-T cells. Sublethally irradiated Rag2-deficient mice were reconstituted with FL-derived pro-B cells and 4 weeks later with a 3:1 mixture of in vitro propagated pro-T cells [22] and Foxp3-transduced in vitro propagated pro-T cells. Mice were immunized with NIP-OVA 11 weeks after pro-B cell transfer and splenocytes analyzed by flow cytometry 3-7 weeks later. (A) IgM and IgG anti-NIP responses of reconstituted and immunized B6 Rag2-deficient mice. Sera were collected 1 week before and 2 weeks after immunization. (B) Representative FACS analysis of splenic B- and T-cell populations 10-14 weeks after pro-T cell transfer. (C) Absolute numbers of T and B cells found in the spleens of reconstituted mice. ns: not significant or $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Unpaired student's t test; n = 7-9. Numbers above data points indicate mean ± SEM. Two independent experiments were performed and each symbol represents an individual mouse.

SUPPORTING INFORMATION

Supporting tables

Provided in the CD

Supporting figures

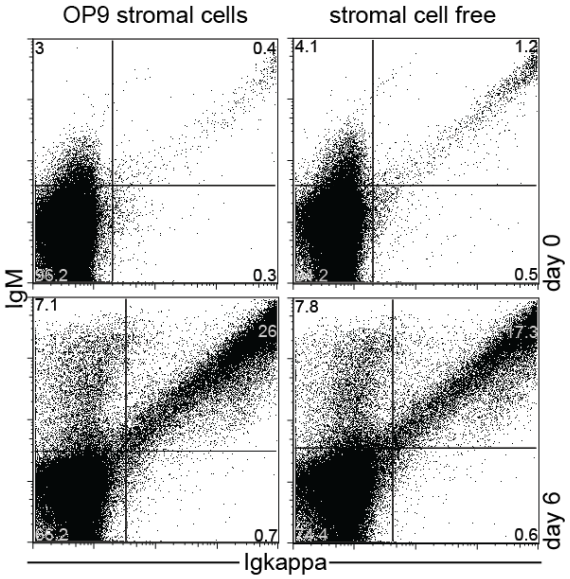


Figure S1
BM EPLM-derived pro-B cells propagated either on stromal cells or grown stromal-cell free differentiated efficiently into IgM positive B cells upon IL-7 withdrawal. Shown is μH and Igκ expression by BM-derived $CD19^{-}B220^{+}CD117^{+}NK1.1^{-}$ cells from Bcl2 transgenic mice cultured on OP9 stromal cells plus IL-7 or cultured in the presence of IL-7, SCF and FLT3L (upper dot blots) and by the same cells cultured for 3 weeks as described above and then for 6 days in plain medium. Whereas in the presence of IL-7 very few cells expressed a μH (IgM) and Igκ light chain on their surface, withdrawal of IL-7 resulted in 26% of pro-B cells cultured on stromal cells becoming IgM^{+}, Igk^{+}. The remaining 7.1% IgM^{+}, Igk^{-} (upper left quadrant) cells being $Igλ^{+}$ (data not shown). These values were 17.3% and 7.8% respectively for pro-B cells cultured in stromal cell-free conditions (lower right dot plot). This has been observed in more than 3 independent experiments.

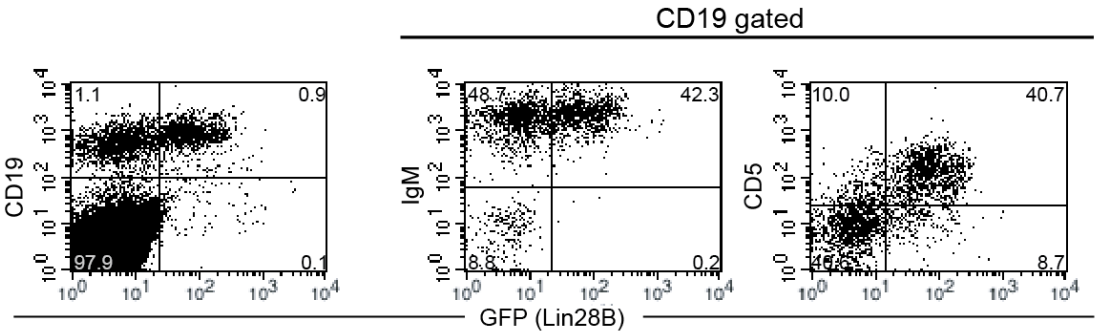


Figure S2
CD19 and GFP expression and IgM, GFP and CD5 expression on CD19-gated cells recovered from B6 Rag2-deficient mice reconstituted with 5 x 10^6 BM-derived pro-B cells and 5 x 10^6 BM-derived Lin28B-IRES-GFP transduced pro-B cells. Analysis was performed 8 weeks after cell transfer. This has been observed in more than three independent experiments.

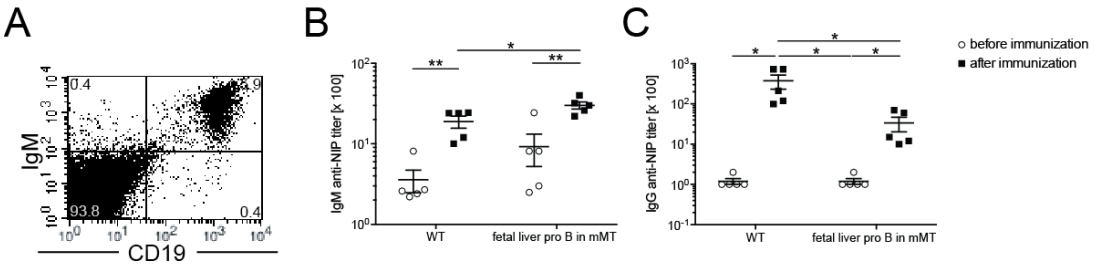


Figure S3
Reconstitution of the B-cell compartment in µMt mice with FL-derived pro-B cells. (A) CD19 and IgM expression by µMT spleen cells 8 weeks after cell transfer. (B) IgM anti-NIP titer of µMT mice immunized with NIP-OVA at 8-10 weeks after pro-B cell transfer. Anti-NIP titers were determined in the sera taken 2 weeks after immunization. Sera taken 1 week before immunization were used as negative controls. Sera from WT B6 mice immunized with NIP-OVA were used as positive controls. (C) IgG anti-NIP titer of µMT mice immunized with NIP-OVA at 8-10 weeks after pro-B cell transfer. Anti-NIP titers were determined in the sera taken 2 weeks after immunization. Sera taken 1 week before immunization were used as negative controls. Sera from WT B6 mice immunized with NIP-OVA were used as positive controls. ns: not significant or $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. Student's t test; n = 5. Data shown above are mean ± SEM.

9.3 Review:

Versatility of stem and progenitor cells and the instructive actions of cytokines shape haematopoiesis

Geoffrey Brown[1], Ciaran J Mooney[1], Llucia Albertí-Servera[2], Lilly von Münchow[2], Kai-Michael Toellner[1], Rhodri Ceredig[3] and Antonius G Rolink[2]

[1] School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[2] Developmental and Molecular Immunology, Department of Biomedicine, University of Basel, Basel, Switzerland

[3] Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland

Critical Reviews in Clinical Laboratory Sciences

http://informahealthcare.com/lab
ISSN: 1040-8363 (print), 1549-781X (electronic)

Crit Rev Clin Lab Sci, 2015; 52(4): 168–179
 DOI: 10.3109/10408363.2015.1021412

informa healthcare

REVIEW ARTICLE

Versatility of stem and progenitor cells and the instructive actions of cytokines on hematopoiesis

Geoffrey Brown[1], Ciaran James Mooney[1], Llucia Alberti-Servera[2], Lilly von Muenchow[2], Kai-Michael Toellner[1], Rhodri Ceredig[3], and Antonius Rolink[2]

[1]*School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham, UK,* [2]*Developmental and Molecular Immunology Laboratory, Department of Biomedicine, University of Basel, Basel, Switzerland, and* [3]*Regenerative Medicine Institute, National Centre for Biomedical Engineering Science, Department of Physiology, National University of Ireland, Galway, Ireland*

Abstract

For many years, developing hematopoietic cells have been strictly compartmentalized into a rare population of multi-potent self-renewing hematopoietic stem cells (HSC), multi-potent hematopoietic progenitor cells (MPP) that are undergoing commitment to particular lineage fates, and recognizable precursor cells that mature towards functional blood and immune cells. A single route to each end-cell type is prescribed in the "classical" model for the architecture of hematopoiesis. Recent findings have led to the viewpoint that HSCs and MPPs are more versatile than previously thought. Underlying this are multiple routes to a particular fate and cells having clandestine fate options even when they have progressed some way along a pathway. The primary role of cytokines during hematopoiesis has long been seen to be regulation of the survival and proliferation of developing hematopoietic cells. Some cytokines now clearly have instructive actions on cell-fate decisions. All this leads to a new way of viewing hematopoiesis whereby versatile HSC and MPP are directed towards lineage outcomes via cytokine regulated cell-fate decisions. This means greater flexibility to the shaping of hematopoiesis.

Abbreviations: **ALL**: acute lymphoblastic leukemia; **AML**: acute myeloid leukemia; **CIC**: cancer-initiating cells; **CLP**: common lymphoid progenitor; **CMP**: common myeloid progenitor; **CSC**: cancer stem cell; **DC**: dendritic cell; **DN**: double-negative; **DP**: double-positive; **EBF**: early B cell factor; **Epo**; erythropoietin; **EPLM**: early progenitors with lymphoid and myeloid potential; **ETP**: early thymocyte progenitors; **Flt3**: Fms-like tyrosine kinase 3 receptor; **Flt3L**: Flt3 ligand; **G-CSF**: granulocyte colony-stimulating factor; **GM-CFU**: granulocyte/macrophage colony-forming unit; **GM-CSF**: granulocyte/macrophage colony-stimulating factor; **GMP**: granulocyte and monocyte progenitor; **HPC**: hematopoietic progenitor cells; **HSC**: hematopoietic stem cell; **IL**: interleukin; **IL-7R**: interleukin-7 receptor; **LIC**: leukemia-initiating cell; **LMPP**: lymphoid-primed multi-potent progenitors; **LSC**: leukemia stem cells; **LSK**: lineage markers , Sca-1$^+$, c-Kit$^+$ population of bone marrow cells; **LT-HSC**: long-term reconstituting hematopoietic stem cell; **M-CSF**: macrophage colony-stimulating factor; **MegE**; megakaryocyte/erythroid; **MEP**: megakaryocyte and erythroid progenitor; **MPP**: multi-potent hematopoietic progenitor cells; **NK**: natural killer cell; **SCF**: stem cell factor; **SLAM**: signaling lymphocyte activation molecule; **SOCS**: suppressor of cytokine signaling; **ST-HSC**: short-term reconstituting hematopoietic stem cell; **TCR**: T cell receptor; **TGF-β1**: transforming growth factor-β1; **Tpo**: thrombopoietin; **TSP**: thymus-settling progenitor; **vWF**: von-Willebrand factor

Keywords

Cell differentiation, cell-fate decisions, cytokines, hematopoiesis, leukemia

History

Received 16 October 2014
Revised 14 January 2015
Accepted 17 February 2015
Published online 27 July 2015

Referees Dr. Isidro Sanchez-Garcia, Instituto de Biología Molecular y Celular del Cáncer, CSIC/Universidad de Salamanca, Salamanca, Spain; Dr. Cesar Cobaleda-Hernandez, Centro de Biología Molecular Severo Ochoa, CSIC/Universidad Autónoma de Madrid, Madrid, Spain; Dr. Tom Taghon, Department Of Clinical Chemistry, Microbiology And Immunology, Ghent University, Ghent, Belgium.

Address for correspondence: Dr Geoffrey Brown, School of Immunity and Infection, College of Medical and Dental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. E-mail: g.brown@bham.ac.uk

Introduction

The hematopoietic stem cell (HSC) gives rise to a wide range of blood and immune cell types. The generation of large numbers of each cell type is a complex and tightly regulated process that is ultimately governed by commitment of rare, and generally quiescent, HSCs to pathways of cell differentiation. These cells, which reside within the bone marrow in adult mammals, are the apex of the hematopoietic hierarchy. HSCs, which can self-renew, give rise to multi-potent

progenitor cells (MPPs) that undergo decision-making, expansion, and differentiation, via recognizable lineage precursors, to give rise to the final compartment of functional cells. Since the early 1980s, the end-cell types have been viewed as two families: lymphoid that includes B and T lymphocytes and natural killer (NK) cells, and myeloid consisting of the rest of the blood and immune cells. Commensurate with this has been the identification of progenitors of each family; namely the common lymphoid progenitor (CLP)[1] and common myeloid progenitor (CMP)[2]. This lymphoid/myeloid dichotomy is the basis of a long-standing model for the architecture of hematopoiesis which also encompasses single routes of differentiation towards individual fates[3]. Over the past 15 years, the above strict architecture has become less black and white. HSCs and hematopoietic progenitor cells (HPCs) are now viewed as more versatile and developing HSCs/HPCs as less rigorously compartmentalized.

Cell heterogeneity and versatility during hematopoiesis

Heterogeneity of hematopoietic stem cells

HSCs are better described in mouse than in human, and mouse HSCs can be purified to homogeneity to a greater extent[4]. There are differences between mouse and human HSCs, and between fetal and adult HSCs. Human HSCs are isolated for transplantation on the basis of expression of the cell surface molecule CD34. In contrast, a single mouse $CD34^{low/-}$ HSC reconstitutes hematopoiesis long-term in a lethally irradiated mouse[5,6]. Matsuoka et al. observed that HSCs of bone marrow, liver, and spleen from fetal and neonatal mice express CD34, whereas HSCs are enriched in the $CD34^-$ fraction and MPPs are $CD34^+$ in mice older than 10 weeks[7]. Albeit, the principles of hematopoiesis derived from human and mouse studies, including of fetal and adult cells, are substantially similar.

The description of HSCs, and isolation by fluorescence activated cell sorting, relies on the presence and absence of a range of cell surface molecules. Studies of HSCs in the mouse have the advantage that HSCs can be rigorously defined during purification as cells that repopulate long-term (LT-HSC) and short-term (ST-HSC) the entire hematopoietic system. A cardinal aspect of HSCs is that they do not express markers that are associated with the various hematopoietic cell lineages (Lin^-), including, for example, CD3 (T lymphocytes), B220 (B lymphocytes), CD11b (monocytes/macrophages), Ly-6G (neutrophils), and TER-119 (erythroid cells). HSCs and MPPs express the two molecules, c-Kit, a mast/stem cell growth factor receptor and tyrosine kinase, and Sca-1, a phosphatidylinositol-anchored membrane protein. LT-HSC and ST-HSC reside in the Lin^-, $Sca\text{-}1^+$, $c\text{-}Kit^+$ population of bone marrow cells, termed LSK.

Originally the Weissman group sub-divided the LSK compartment on the basis of expression of Thy-1.1 and the Flt3 (fms-like) tyrosine kinase, a type 3 receptor kinase[3,8]. The loss of a low level of expression of Thy-1.1 and gain of expression of Flt3 was observed to correlate with a loss of the capacity of HSCs to self-renew (Figure 1). LT-HSCs were isolated from adult bone marrow as LSK $Flt3^-$, and

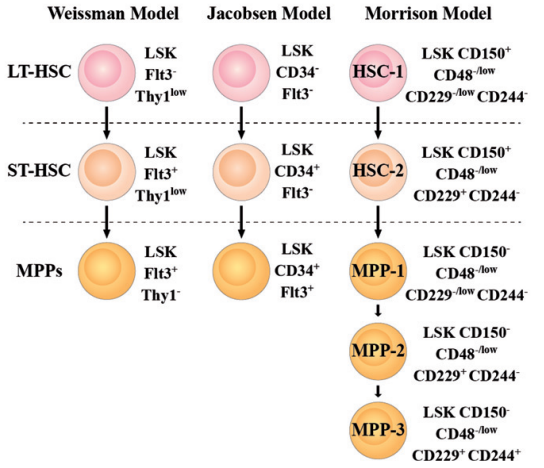


Figure 1. Identification of hematopoietic stem and progenitor cells in adult bone marrow. The loss and gain of cell surface molecules is used to delineate sub-populations of hematopoietic stem cells (HSCs) and multi-potent progenitors (MPPs). Weissman subdivides the lineage markers , $Sca\text{-}1^+$, $c\text{-}Kit^+$ (LSK) population of bone marrow cells on the basis of expression of Flt3 and Thy1.1. Jacobsen makes use of CD34 and Flt3 expression. Signaling lymphocyte activation molecule (CD150, CD48, CD229, and CD244) markers are used to subdivide LSK cells in the model proposed by Morrison. These markers are representative for hematopoiesis in mice. Flt3, fms-like tyrosine kinase; LT-HSC, long-term reconstituting hematopoietic stem cells; ST-HSC, short-term reconstituting hematopoietic stem cells.

transplantation of LSK $Flt3^+$ cells resulted in short-term multi-lineage reconstitution. However, fetal liver HSCs are contained within the $Flt3^+$ and $Flt3^-$ LSK cells. MPPs are $Thy1^-Flt3^+$. In 2005, Yang et al. combined LSK markers with CD34 and Flt3. The definitions provided by the Jacobsen group, which are commonly used, are LT-HSC as LSK $CD34^-Flt3^-$, ST-HSC as LSK $CD34^+$ $Flt3^-$, and MPPs as LSK $CD34^+$ $Flt3^+$ (Figure 1)[9].

Recently, the signaling lymphocyte activation molecule (SLAM) family of membrane receptors (CD150, CD48, CD229, and CD244) has been used to describe subpopulations of HSCs and MPPs (Figure 1). Kiel et al. purified HSCs and MPPs as $CD150^+CD48^-CD244^-$ and $CD150^{-/+}CD48^-CD244^+$, respectively. Lineage-restricted progenitors can be identified by acquired expression of CD48 and are mostly $CD150^-CD48^+CD244^+$[10]. Adding CD229 to the panel of SLAM, CD34 and LSK markers led Oguro et al. to propose HSC-1 and HSC-2 populations, which have different lineage biases, and three populations of MPPs (Figure 1). $CD229^-$ HSC-1 are cells that rarely divide and are myeloid-biased as revealed by transplantation studies. $CD229^+$ HSC-2 cells divide more frequently and are lymphoid biased[10]. Contrary to the findings of the Morrison group, the level of expression of CD150 has also been used to distinguish myeloid- and lymphoid-biased HSCs. Ema et al. have observed that myeloid-biased HSCs are enriched in the $CD150^{high/med}CD34^-LSK$ population and lymphoid-biased HSCs are enriched in $CD150^{low/-}CD34^-LSK$ population. These workers have proposed that these two populations of cells overlap with LT-HSCs and ST-HSCs[11].

As mentioned above and for many years, the two compartmentalizing properties of HSCs are their capacities to reconstitute the entire hematopoietic system and to self-renew. Oguro, in summarizing the sub-populations of hematopoietic stem and progenitor cells distinguished using SLAM markers, provides a scheme whereby the self-renewal potential of HSC-1 is long-term, of HSC-2 is long to intermediate, of MPP-1 is intermediate to transient, and of MPP2 and MPP3 is transient[10]. Oguro also described HPCs (HPC-1 and 2) which do not self-renew. In essence, the boundary between HSCs and MPPs as to the property of self-renewal is blurred, and a distinction between HSCs and MPPs is perhaps somewhat redundant. Both these cell types might be best viewed, and classified as a continuum of HPCs that reduce their capacity to self-renew as they mature.

Strict multi-potency does not always go hand in hand with self-renewal as to the identification of HSCs with lineage biases. As mentioned above, the CD34$^-$LSK population of bone marrow cells, which engraft mice including secondary hosts, has been divided into myeloid-biased and lymphoid-biased HSCs. In addition to expressing a higher level of CD150, myeloid-biased HSCs exclude Hoechst 33342 more effectively than lymphoid-biased HSC, and these two HSC sub-types are differentially regulated by transforming growth factor-β1 (TGF-β1)[12]. There is additional heterogeneity within the CD150^{high}CD34$^-$LSK population of cells, as revealed by engraftment of single cells in irradiated mice. In primary hosts, some single cells readily gave rise to myeloid cells and cells that were able to engraft a secondary host. Some single cells produced few myeloid cells in primary hosts and cells which when transferred to secondary hosts gave rise progressively to multiple lineages[13]. Furthermore, CD41 and CD86 expression on HSCs has been reported to distinguish myeloid-biased and lymphoid-biased cells, respectively[14,15]. As mice age, there are quantitative differences in lineage biases in the HSC population. The data support a model whereby myeloid-biased HSCs have clonally expanded, while lymphoid-biased HSCs exhaust themselves due to their more extensive proliferative nature[16].

Recently Jacobsen's group have described a platelet-biased LT-HSC that expresses von-Willebrand factor (vWF)[17]. Transplantation of single vWF$^+$ HSCs into irradiated hosts resulted in reconstitution biased towards platelets and myeloid cells. vWF$^-$ HSCs gave rise to a lymphoid-biased reconstitution. vWF$^+$ HSCs require thrombopoietin (Tpo) for their maintenance as these cells were significantly reduced in number in Tpo$^{-/-}$ mice. vWF$^+$ HSCs gave rise to vWF$^-$ HSCs, and vWF$^-$ HSCs were not able to give rise to vWF$^+$ HSCs. This led the Jacobsen group to propose that platelet-biased HSCs are the apex of the hematopoietic hierarchy.

Heterogeneity of hematopoietic progenitor cells

In the ''classical'' lymphoid/myeloid dichotomy model of hematopoiesis, the sets of potentials observed for different types of progenitors align themselves to developmental progression along each arm of the dichotomy. However, this is not the case for all the progenitors that have been described to date. At odds with an irrevocable commitment of HSCs to either a lymphoid or myeloid pathway is the early description

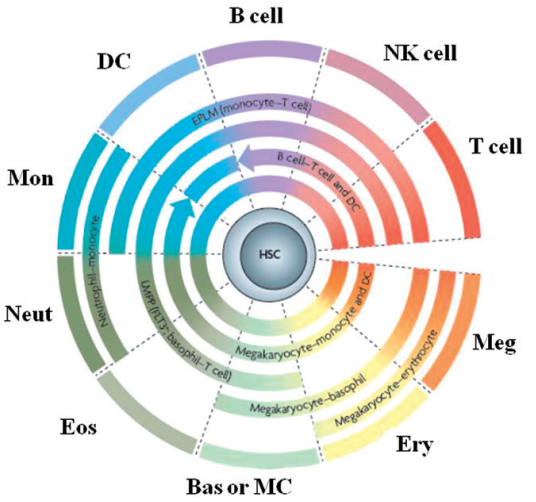


Figure 2. A pair-wise model of hematopoiesis. A fate choice continuum with an invariant series of pair-wise developmental relationships between hematopoietic cells fates is derived from the nature of the sets of potentials of various hematopoietic progenitor cells[22–25]. These are show as segments of the continuum. Dendritic cells (DC) can be derived from common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs), as shown by arrowheads on two of the arcs[26]. The figure, modified and with permission is from Ref. (21) © Macmillan Magazines Ltd. Bas, basophil; Eos, eosinophil; EPLM, early progenitor with lymphoid and myeloid potential; Ery, erythroid; MC, mast cell; Meg, megakaryocyte; Mon, monocyte; Neut, neutrophil; NK, natural killer cell.

of a progenitor in mouse fetal liver with just the potentials for B lymphoid and macrophage differentiation[18]. This cell was later shown to be present in adult bone marrow[19]. A further finding that contradicts the notion that HSCs make an immediate and irrevocable decision to commit to either the lymphoid or myeloid pathways of differentiation is the identification of cells with lymphoid potentials and an incomplete set of myeloid potentials. These cells are early progenitors with lymphoid and myeloid potential (EPLM), that can give rise to T and B lymphocytes, NK cells, dendritic cells (DCs), and macrophages[20], and lymphoid-primed multipotent progenitors (LMPP), that have little potential for megakaryocyte or erythroid development while retaining other potentials[21].

Progenitor cells that contradict a lymphoid/myeloid dichotomy led us to propose the pair-wise model of hematopoiesis[22–24]. This model does not assume lineage branching patterns or prescribe a single preferred route to a particular end-cell fate. Instead mature cell fates are shown to be near-neighbors within a continuum of lineage fates (Figure 2). As HSCs mature towards a specific dominant cell fate, fates that are distantly related to this fate are lost first and more closely, related fates remain possible as latent fates. The model envisages versatility of HSCs and MPPs, as to allowing an end-cell type to be reached by more than one route (see also below). Mapping of transcription factor usage and the responsiveness of progenitors to growth factors support the proposed close relationships between cell lineages[22–24].

DOI: 10.3109/10408363.2015.1021412

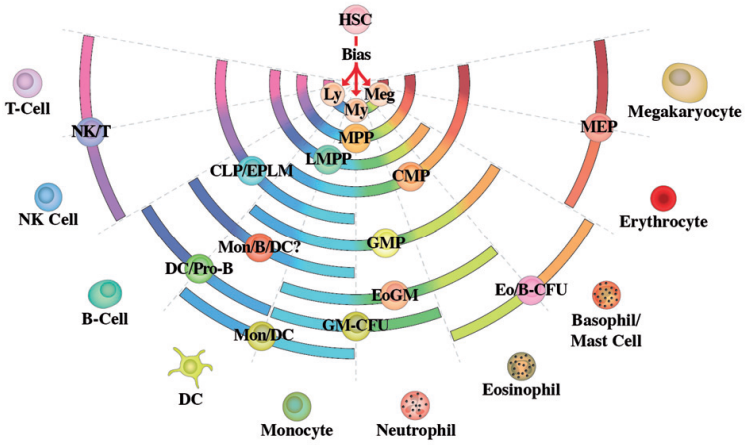


Figure 3. Mapping of hematopoietic progenitor cells to the pair-wise model. The set of fates available to various progenitors underlie near-neighbor placement of lineages in the pair-wise model. Progenitors with commensurate and contiguous lineage potentials are CLP, common lymphoid progenitor[1]; CMP, common myeloid progenitor[2]; DC/Pro-B, dendritic cell and B lymphocyte progenitor[26]; Eo/B-CFU, eosinophil and basophil progenitor[28]; EPLM, early progenitor with lymphoid and myeloid potential[20]; GMP, granulocyte and macrophage progenitor[27]; LMPP, lymphoid-primed multi-potent progenitor[21]; MEP, megakaryocyte and erythrocyte progenitor[29]; Mon/B/DC?, monocyte, B lymphocyte and dendritic cell? progenitor[18,19]; Mon/DC, monocyte and dendritic cell progenitor[30]; NK/T, natural killer cell and T lymphocyte progenitor[31]; HSC, hematopoietic stem cell; Ly, lymphoid bias; Meg, megakaryocyte bias; My, myeloid bias.

A number of progenitor cells that have different sets of potentials have been described. The various combinations of lineage potentials that exist within progenitors are also reflected in a number of different cell lines[25]. Figure 3 shows that the different combinations of differentiation capabilities described for normal progenitors can be mapped to the pair-wise model. Of particular importance to placing cell lineages close to one another in the continuum are progenitor cells that have just two differentiation capacities. In other words, the continuum infers that only certain bi-potentialities are permissible. For example, a cell with the potential for megakaryocyte and T cell differentiation should not exist, and has not been described to date. Bi-potent progenitor cells can be placed within the model with the exception of a bi-potent B lymphocyte/macrophage cell[18,19]. Whether this cell can give rise to DCs has not been studied.

Already there is a considerable variety of stem cells, with differing biases, and progenitor cells with differing sets of potentials. The latter in turn give rise to end-cell types that can be divided into numerous sub-types, for example, as is the case for T helper cells and DCs. The full extent of the heterogeneity of progenitor cell populations, and their mature progeny, is very much contingent on the extent to which new and existing cell surface markers can be used to define new sub-populations. It is highly likely that progenitor cells with multiple lineage options and that we presently view as a homogeneous population of cells will be divided into cells with lineage biases in various directions.

Versatility of hematopoietic progenitor cells

The pair-wise model allows developmental pathways to be flexible. One aspect of this is HSCs and their progeny using more than one route to a particular end-cell type. This principle was demonstrated by deriving DCs *ex vivo* from cells purified as CLPs and CMPs. When the transcription profiles of the two DC populations were compared they were found to be the same[26]. Ishikawa et al. concluded that the developmental program of human DCs operates independently of the pathways for myeloid and lymphoid cells.

Alongside new findings there has been a plethora of new models for the architecture of hematopoiesis. Some of the models depict multiple routes towards granulocytes and monocytes. For example, a model provided by Jacobsen considers the possibilities of these myeloid cells arising from ST-HSC/MPP via (i) CMP and GMP (granulocyte and monocyte progenitor); (ii) LMPP giving rise to GMP; and (iii) LMPP giving rise to a granulocyte/monocyte/T lymphocyte progenitor[32] which in turn gives rise to GMP. In a model proposed by Katsura, the two routes towards myeloid cells are HSCs veering towards: (i) a cell with the potentials for myeloid, erythroid, and megakaryocyte development; and (ii) a cell with the potentials for myeloid and lymphoid development[33]. Ye and Graf compared the production of mature cell types to flows along branches of a tree with a major branch giving rise to platelets, erythroid cells, granulocytes, and monocytes in equal measure and a separate branch giving rise largely to granulocytes and monocytes[34]. Of course, it is difficult to exclude the possibility that all the above routes to myeloid cells occur to some degree.

Precise tracking of a progenitor, as defined by a set of markers, giving rise to the next progenitor, also verified by markers, and so on to an end-cell type(s) is an impossible task. However, a novel approach has been recently described whereby cells can be uniquely marked *in situ*, by transposon tagging, allowing longitudinal analysis of clones in mice and revealing that long-lived progenitors are a main driver of steady-state hematopoiesis in adulthood[35]. This approach will also impact on our view of cell lineages and provide information on the origin of hematopoietic malignancies. Aside from this experimental model, examination of the sets of options available to various progenitors, as defined by appropriate markers, and which progenitors are or are not able via loss of a fate option(s) to give rise to one another, allows configuration of routes to end-cell types. For certain cell types, it is clear that there can be multiple routes. As described by Jacobsen et al., HSCs can give rise to neutrophils and monocytes through LMPP and CMP intermediates. Figure 4 shows different possible routes downstream of LMPP and CMP towards neutrophils. These include

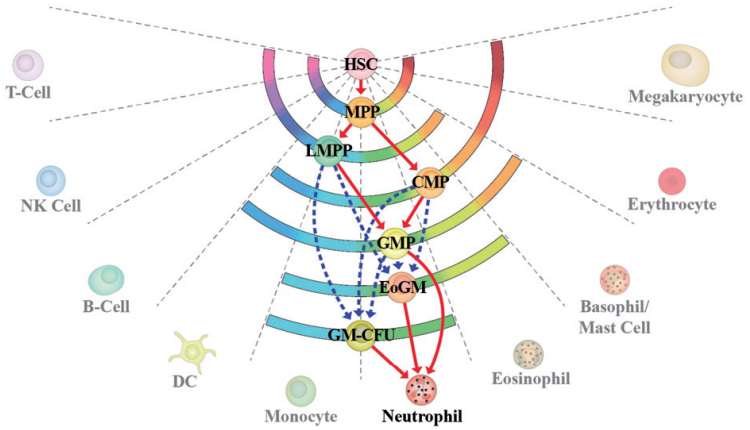


Figure 4. Alternative developmental routes towards neutrophils. The sets of potentials available to known oligopotent progenitors are used to construct possible routes. The red solid arrows are routes delineated from studies of the progeny of progenitors and the blue dash arrows are putative routes. For abbreviations, see the legend to Figure 3.

potential pathways through an eosinophil/granulocyte/monocyte progenitor[36] and/or GM-CFU (granulocyte/macrophage colony-forming unit). The routes are in keeping with the pairwise model as to their close proximity. Similarly, there are multiple potential pathways towards monocytes via CMP, an eosinophil/granulocyte/monocyte progenitor and GM-CFU, and *via* LMPP, EPLM, and a monocyte/dendritic cell progenitor[30]. In fetal liver, a myeloid/B-biased progenitor is a further intermediary to myeloid cells[37]. A caveat to all this is the extent to which routes are used *in vivo*. However, the point of interest is the inherent flexibility of pathways available to progenitors. It is noteworthy that the ability of precursor cells to follow alternative developmental pathways to give rise to the same cell phenotype was described for cell lineages in the embryo of the leech as early as 1987[38].

One aspect underpinning versatility is that progenitor cells retain fate options even after they have progressed some way along a pathway. Thymus-settling progenitors (TSP) have the potential to give rise to myeloid cells, dendritic cells, NK cells, and B lymphocytes in addition to T lymphocytes. TSP give rise to double-negative (DN) 1 early thymocyte progenitors (ETP), which give rise to DN2 cells. DN2 cells have lost the potential for differentiation towards B lymphocytes, but when cultured in the right environment generate myeloid, dendritic cells, and NK cells[39–41]. As to driving pro-T cells to other fates, culture of DN2 cells in IL3 (interleukin) and stem cell factor (SCF) in the absence of Notch signaling revealed mast cell potential, indicating a close unexpected relationship between the pro-T cell and mast cell fates[42]. These clandestine potentials are lost as DN2 cells progress to the DN3 stage of development.

Some cytokines have instructive actions on cell-fate decisions

Early glimpses to the instructive action of cytokines

A long-standing debate is whether the commitment of HSCs to fate options occurs in a cell-autonomous and stochastic way, or is driven (in an ordered way) by instructive signals from the local environment[43–46]. Cytokines are the pivotal external factors that impart environmental signals to control hematopoietic cell development. They have multiple actions that can be viewed as either instructive, by directing HSCs/MPPs towards a specific lineage, or permissive, by selectively allowing cells committed to a particular lineage to survive and/or proliferate[47,48]. For many years, a permissive role of cytokines has been favored. A very recent and complete turnabout in our understanding of the control of hematopoiesis is the notion that cytokines instruct decision-making[49,50].

Information to support an instructive role for cytokines has been available for quite some time. In 1982, Metcalf and Burgess concluded that granulocyte/macrophage colony-stimulating (GM-CSF) factor and macrophage colony-stimulating factor (M-CSF) can ''*irreversibly commit the progeny of GM-CFC respectively to granulocyte and macrophage production*''[51]. When paired daughter cells of GM-CFU were split and one cultured in GM-CSF and the other in M-CSF, some of the cells underwent irreversible commitment to the granulocyte and macrophage pathways, respectively. This occurred during completion of the first cell division and within 24 h. Later in 1991, Metcalf again concluded that colony-stimulating factors have the ability to influence lineage commitment[52]. Metcalf examined the relative frequencies of lineage committed progenitors when blast cell colonies were established from normal bone marrow cells in combinations with granulocyte colony-stimulating factor (G-CSF), GM-CSF, and multi-CSF. The relative frequency of granulocyte progenitors was increased when cultures were established in the combination with GM-CSF or multi-CSF with SCF.

In 2000, Kondo et al. provided more evidence to support the notion that cytokines can convert the fate of lymphoid-committed progenitors[49]. The IL-2 and GM-CSF receptors were exogenously expressed in CLPs, which normally gives rise exclusively to T lymphocytes, B lymphocytes, and NK cells. This resulted in cell-fate conversion to the myeloid lineage. The use of mutants of the beta-chain of the IL-2 receptor revealed that signals for the granulocyte and monocyte differentiation pathways are provoked by different cytoplasmic domains of the IL-2 receptor. Kondo et al. also showed that primitive HSCs express low-to-moderate levels

of the receptors for GM-CSF and M-CSF. Hence, there is a possibility that HSCs are receptive to the instructive actions of these growth factors. Kondo et al. concluded from all the above that down-regulation of cytokine receptors that drive myeloid cell development is a critical step in commitment of cells to lymphoid development.

Recent studies confirm an instructive action of cytokines

A number of studies have now shown that cytokines have an instructive action on cell-fate decisions. M-CSF, G-CSF, and erythropoietin (Epo) instruct monocytic, neutrophilic and erythroid fates, respectively. In 2009, Rieger et al., by monitoring individual hematopoietic progenitors in culture, confirmed that M-CSF and G-CSF provided instructive cues. Using bioimaging techniques and a LysM-GFP reporter system to detect differentiation and cell death, the group observed that individual GMPs adopted monocytic or neutrophilic fates in the presence of M-CSF or G-CSF, respectively[50]. More recently, Mossadegh-Keller et al. made use of PU.1-GFP reporter mice to show that M-CSF drives expression of myeloid-associated genes in some LT-HSCs. *In vivo*, M-CSF stimulated expression of PU.1 HSCs generated increased numbers of GMPs in the spleen and peripheral myeloid cells at the expense of cells undergoing megakaryocyte, erythroid, and lymphoid development when compared with non-primed PU.1$^-$ HSCs[53]. Epo induces priming of erythroid lineage-associated genes in LT-HSCs and *in vivo* skews the potential of these cells towards an erythroid fate[54]. An increase in serum Epo levels in mice led to the expansion of committed erythroid and megakayocyte and erythroid progenitors (MEP) in the bone marrow, whereas megakaryocyte progenitors, pre-GMP, and LMPP populations were decreased.

An instructive action of Flt3 ligand in determining cell fate

One of the cytokines that is essential to cell survival and proliferation during early hematopoiesis is the ligand for Flt3[55,56]. Flt3 ligand (Flt3L) was described two decades ago[55] and is the only known ligand for Flt3. Upon ligand binding, the Flt3 receptor dimerizes and initiates signaling that involve STAT5a, ERK1/2, and PI3K[57]. Flt3 is an important area of research since mutations in Flt3 were among the first ones discovered in acute myeloid leukemia (AML)[58]. Presently, there is substantial interest in elucidating the instructive role of Flt3L.

Flt3 expression occurs during hematopoiesis at the non-self-renewing ST-HSC stage of development[59]. In fact, Flt3 up-regulation relates to the loss of self-renewal capacity[60]. This may just be a coincidence, or Flt3/Flt3L provokes the loss of self-renewal capacity by an as-yet-unknown mechanism. Thereafter, MPP express Flt3 as do several downstream progenitors with myeloid and/or lymphoid potential, while the MEP is Flt3$^-$[59]. As lineage options become more restricted, Flt3 expression is down-regulated with the exception of DCs[61].

Importantly, Flt3L exerts a role by interacting with other cytokines such as IL-7 or SCF[62,63]. For example, IL-7 and

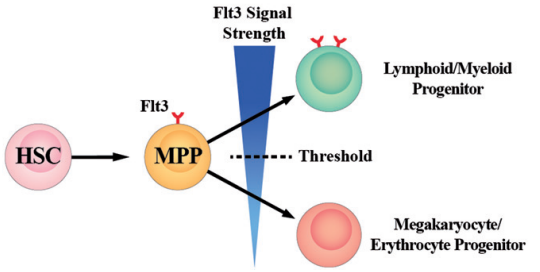


Figure 5. Instructive action of Flt3 ligand in determining lymphoid/myeloid *versus* MegE lineage development. If the Flt3 ligand signal strength exceeds a certain threshold level, cells enter the lymphoid/myeloid branch at the expense of the MegE lineage. HSC, hematopoietic stem cell; MPP, multi-potent progenitor.

Flt3L stimulate lymphoid development in a coordinated manner that occurs in a narrow window during which cells express both receptors. In the case of *in vitro* cultures of hematopoietic progenitor cells, such as ETP, CLP, or EPLM, and in the presence of IL-7, Flt3L provides an additive anti-apoptotic effect while stimulating proliferation[64,65] (and our observations). This additive effect is reported to result from parallel activation of the IL-7 receptor (IL-7R) and Flt3 via separate signaling pathways converging to activate Stat5[63].

Mice with targeted gene disruption of Flt3[66] or its ligand[67] have provided precise information about the actions of Flt3/Flt3L. These mice have defects in the developmental potential of myeloid/lymphoid progenitors as well as reduced numbers of B cells, DCs, and NK cells[66,67]. Tsapogas et al. have provided evidence to support the notion that Flt3L is instructive to cell decision-making. These workers generated a Flt3L transgenic (Flt3L-Tg) mouse that expresses human Flt3L[68]. Flt3L-Tg mice have a tremendous expansion of hematopoietic progenitors in the bone marrow. Of all the progenitor populations analyzed, the only progenitor that was decreased was the MEP. As such, the Flt3-Tg mice had decreased platelet counts and developed anemia. Previous studies had reported that the expression of Flt3 after the ST-HSC stage leads to a reduction of megakaryocyte and erythrocyte potentials[21]. Also, a significant reduction in MEP progenitors was observed by day 3 when wild-type mice were injected with recombinant Flt3L[69]. Considering this rapid response, it is likely that the reduction in MEP numbers was a consequence of a Flt3L threshold response in upstream Flt3$^+$ progenitors rather than a space restriction within the bone marrow caused by the over-proliferation of other progenitors. It was proposed that upstream Flt3$^+$ progenitors develop towards lymphoid/myeloid lineages on receiving Flt3L stimulation above a certain threshold level and develop towards the MegE (megakaryocyte/erythroid) lineage if this threshold level was not reached. In other words, an increased level of Flt3L guides the development of cells towards the lymphoid/myeloid fates at the expense of the MegE fates (Figure 5). This provides an explanation of the MegE developmental defect in the Flt3L-Tg mice. The exact mechanism by which Flt3L exerts an instructive action remains to be elucidated and whether this occurs at the CLP/EPLM level is of considerable interest.

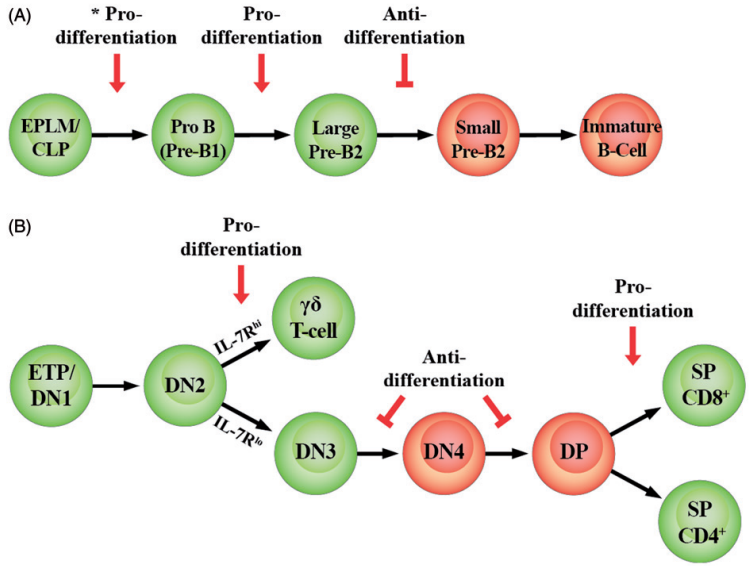


Figure 6. IL-7 action during various stages of murine B and T lymphocyte development. The progressive stages of B lymphocyte development (A) and T lymphocyte development (B) are shown. EPLM/CLP to Large Pre-B2 (green cells in A) and ETP/DN1 to either γδ T-cell or DN3 and SP (green cells in B) represent stages responsive to the survival and proliferative action of IL-7. Pro-differentiation and anti-differentiation actions are also shown.*The IL-7 pro-differentiation action could be instructive. For other stage transitions, more evidence is needed to conclude an instructive action. DN, double-negative; DP, double-positive; EPLM, early progenitor with lymphoid and myeloid potentials; ETP, early thymocyte progenitor.

Cytokines play different roles at various developmental stages

This is best illustrated by consideration of the actions of IL-7 at various stages of B and T lymphocyte development[70–72]. These include promoting cell survival and proliferation and facilitating decision-making during differentiation (Figure 6). IL-7 was described in 1988[73]. In mice, deficiencies in IL-7 or the receptor lead to impairment of B and T lymphopoiesis[74–76]. IL-7 is not required for B lymphopoiesis in humans, as deficiencies in IL-7 result in apparently normal B cell development[72]. However, the B cells could be the result of fetal development since they can be cultured *in vitro* in the absence of IL-7 while postnatal B cells seem to be IL-7 dependent[77].

CLPs express IL-7R, which provides a criterion used to isolate these cells[1]. CLPs are dramatically reduced in IL-7-deficient mice[78], indicating that this cell is critically dependent on IL-7. EPLMs also express IL-7R[20]. A strong proliferative action of IL-7 is seen when CLPs and EPLMs, sorted from mouse bone marrow, are cultured on OP9 stromal cells and treated with IL-7[78] (and our own observations). Furthermore, these cells are able to differentiate to the next B cell developmental stage, the pro-B (pre-B1), indicating that IL-7 acts as a differentiation factor at the CLP/EPLM level. In keeping with this notion is that CLP cells are unable to reach the pro-B (pre-B1) stage in $\gamma c^{-/-}$, IL-7R$\alpha^{-/-}$, or IL-$7^{-/-}$ mice, and instead arrest at an uncommitted level[79]. These data suggest an instructive role for IL-7. This appears to occur via IL-7-induced STAT5 signaling which regulates the expression levels of EBF (early B cell factor)[78,80]. In turn, EBF activates transcription of Pax5[81,82] leading to the expression of CD19. This cell surface marker indicates that cells have gained the pro-B phenotype and commitment to the B cell lineage[83,84].

The pro-B (pre-B1) stage is also sensitive to the action of IL-7. Fetal and bone marrow pro-B (pre-B1) cells can be cultured *in vitro* in the presence of IL-7 and stromal cells for long periods of time (more than 4 months)[85]. The survival action of IL-7 appears to be exerted *via* expression of the gene encoding the anti-apoptotic protein Mcl-1[86]. Moreover, there is evidence to suggest that IL-7 facilitates differentiation of pro-B (pre-B1) cells into cytoplasmic μ immunoglobulin (Ig) heavy chain expressing large pre-B2 cells[87,88] (and our own unpublished observation). IL-7 is not instructive, since cells differentiate irrespective of the presence of IL-7[87,88]. Instead, the data argue for a permissive role of IL-7 in promoting survival and proliferation during the pro-B (pre-B1)/large pre-B2 transition. However, some groups have reported that STAT5-mediated IL-7 signaling controls chromatin accessibility and the rearrangement of distal V_H genes at the *Igh* locus[89–91]. This is exemplified by a significant decrease in distal V_H rearrangements in $B220^+$ IgM^- bone marrow B lymphocytes of $IL7R^{-/-}$ and *Stat5*$^{-/-}$-deficient mice[89,91]. In contrast, Malin has argued that there is no substantial difference in the distal V_H rearrangement genes seen for STAT5 deficient, IL-7R mutant, and control pro-B (pre-B1) cells[86]. Finally, as to the pre-B1/pre-B2 transition, it has been reported that IL-7 acts to prevent premature rearrangements of the Igκ via binding of IL7-mediated STAT5 to the Igκ intronic enhancer (iEκ)[86,92].

The large pre-B2 transition is the last stage in B cell development that is sensitive to the action of IL-7. The proliferation of these cells is improved by the presence of IL-7[93]. However, IL-7 acts as an anti-differentiation factor towards immature B cells by blocking the rearrangement of the light chain loci[86,92,93]. Upon withdrawal of IL-7 from the culture, cells undergo Igκ recombination and differentiation into IgM positive cells[21,86,94,95]. At this stage, as a consequence of the pre-B cell receptor signaling, IL-7R is

down-regulated and the later stages of B cell development are unresponsive to IL-7[93,96,97].

During T lymphocyte development, the thymus is seeded by ETP[98] and there is a general agreement that these cells can give rise to multiple lineages. In humans, the ETP retain lympho-myeloid potentials and can generate separate lymphoid- and myeloid-primed progenitors. Lymphoid-restricted progenitors seem to be the main thymus colonizers in the mouse[99], although several *in vitro* studies suggest that these cells have myeloid potentials[64,100]. IL-7Rα is a direct target of Notch1[101], the master signaling pathway that regulates thymopoiesis[102–104]. Up-regulation of IL-7R at the ETP stage[105] is a hallmark for lymphoid commitment, and IL-7-mediated signaling triggers the first wave of expansion of lymphoid-primed progenitors[106,107]. IL-7 is essential to survival and proliferation of ETP, and appears to transmit signals for the survival and proliferation of lymphoid-committed cells at the expense of the myeloid branch[72]. Sustained Notch signaling favors cell development along the T cell lineage by inducing the transcription of important T lymphocyte differentiation factors and blocking the development of cells towards other lineages[105,108]. IL-7 is also required for the transition of ETP cells to the DN2 stage *in vitro*[64]. Moreover, an early block is observed in T lymphocyte development in IL-7 or IL-7R-deficient mice[109]. However, Bcl-2 is sufficient to rescue T lymphocyte development (at least the α/β branch of T lymphocytes), thus confirming a permissive role for IL-7 during the early thymopoiesis[110].

IL-7Rα chain expression increases progressively until the DN2 stage, coinciding with the first massive cellular expansion, and then steadily decreases. DN2 cells with high IL-7R expression levels are diverted to the γδ T cell lineage, at least in mice[111], and IL-7R signaling controls accessibility to the T cell receptor (TCR)γ locus and its rearrangement[112]. Thus, IL-7 plays an instructive role in this developmental process and Bcl-2 is unable to rescue the development of γδ T cells[110,113]. The cells with reduced IL-7R expression and/or limited IL-7 availability progress to the DN3 stage. The proliferation of these cells is still IL-7 dependent[110,113]. However, a diminished IL-7R signaling seems necessary for DN2 mouse thymocytes to up-regulate Bcl11b, a transcription factor that is essential to the T cell lineage[114]. Thereafter, on successful rearrangement of the TCRβ chain and subsequent pre-TCR expression, IL7-R is down-regulated, and cells move to the DN4 stage[115,116]. Here, at the β selection checkpoint, there is a second wave of expansion that is controlled by Notch1 signaling and pre-TCR expression[64]. As such, the DN3 to DP (double-positive) transition is IL-7 independent. However, the continued presence of IL-7 in *in vitro* cultures blocks differentiation and DP cells are generated only on IL-7 removal. In this case, IL-7 is acting as an anti-differentiation factor[64,117]. This highlights the importance of the IL-7R signaling suppression during the DP transition, which is guaranteed by SOCS-1 (suppressor of cytokine signaling)[115]. After the DP selection stage, IL-7R surface expression levels are restored and IL-7 signaling seems to be required for the DP to single positive CD8 transition[118]. In keeping with this, Park et al. reported differentiation of single positive CD8 cells ($CD8^+$) and RUNX-3 transcription factor up-regulation in transgenic TCR mice expressing a transgene-derived IL-7Rα chain in an IL-7 over-expressed environment[119]. Finally, IL-7 is required for homeostatic expansion of naïve $CD8^+$ and $CD4^+$ T cells[120,121].

In the context of the role of IL-7 in lymphopoiesis and in particular the flexibility of the hematopoietic process, the effect of pregnancy on hematopoiesis and lymphopoiesis is worthy of mention. Pregnancy is a physiological process in which there is an increase in blood volume and the bone marrow responds by increasing erythropoiesis[121]. Lymphopoiesis is also affected, and two mechanisms account for the reduced B lymphopoiesis observed in pregnancy. First, early B progenitor cells are sensitive to the increased levels of sex steroids in pregnancy[122] and second, in mice, IL-7 production and availability dramatically decrease[123]. Not only is the bone marrow affected during pregnancy but also the maternal thymus undergoes dramatic involution[124]. Both B and T lymphopoieses return to pre-pregnancy levels following parturition and weaning. A full explanation of these associated phenomena is not currently available and is certainly worthy of further investigation. In regard to the effect of cytokines on lymphopoiesis in the developing fetus, an interesting and recent finding is that inflammatory cytokines interferon- (IFN) α and γ positively regulate the number of lymphoid progenitors in mouse embryos, and IFN-γ signaling also affects the number of HSCs[125]. These findings have important implications to enhancement of the formation of hematopoietic stem and progenitor cells in the embryo and the differential regulation of hematopoiesis during fetal and adult life.

Implications of heterogeneity and versatility to leukemia

A Darwinian viewpoint on leukemia

A widely-held viewpoint for a number of years is that the cell population that makes up an overt leukemia and solid tumor is heterogeneous and that there is a hierarchical organization. A small subset of cells, which are at the apex of the hierarchy, is responsible for maintaining the long-term growth of the leukemia and tumor mass. These leukemia- and cancer-maintaining cells, which have an inherent capacity to self-renew, are termed leukemia stem cells (LSC) and cancer stem cells (CSC), respectively[126,127]. There is also the question of the normal cell from which the leukemia/tumor has arisen, which are termed leukemia-initiating cells (LIC) and cancer-initiating cells (CIC), respectively. These are the cells that suffered an oncogenic lesion and that evolve to give rise to the leukemia/tumor. And, many leukemias and cancers arise from a self-renewing stem cell or downstream progenitor if self-renewal is reactivated. The origin and LSC/CSC nature of leukemias/cancers is important to putting into practice a cancer stem cell-based therapeutic to cure patients[128–132].

To add to the difficultly of designing therapies to eliminate LSC/CSC, Greaves has proposed a Darwinian viewpoint on the nature of CSC which pays attention to the dynamics of the cancer[133,134]. Studies of single acute lymphoblastic leukemia (ALL) cells have revealed that the leukemia "stem" cells are genetically diverse. Greaves has likened the complex and branching clonal architecture of ALL "stem" cells to Darwin's evolutionary tree-like divergence diagram which was drawn in 1837. To add to the Darwinian analogy, as the leukemia evolves,

in an almost entirely clinically silent manner, cells acquire gene copy number alterations. The cytokine TGF-β has been proposed to exert a selective Darwinian advantage to expand the cells that are first "hit" as to genomic alteration and at risk of acquiring secondary multiple gene alterations. In essence, this is similar to "natural selection" as described by Darwin.

Often what we discover from studies of cancer cells, and ascribe to these cells, turns out to be a feature of normal cells. A simple example is that the common ALL-associated antigen (CD10) was first described as a candidate leukemia-specific antigen[135]. In fact, CD10 is expressed by rare B cell progenitor cells and its presence on ALL cells was telling us something about the origin of common ALL[136]. So, we might view normal hematopoiesis as a Darwinian process that is driven by selective pressures, namely cytokines, acting on an inherent diversity that is sufficient to generate the various types of blood and immune cells. In essence, inherent cell heterogeneity must have been the template to the evolution of the wide variety of immune cells that exist in higher mammals.

Are leukemia stem cells as versatile as their normal counterpart?

An important question that arises from the above considerations is the following: are LSCs as versatile as normal HSCs in terms of their capacity to access routes to end-cell types? There is evidence to suggest that LSCs are less versatile in this regard[137,138], as exemplified by erythroleukemia [acute myeloid leukemia (AML) FAB-M6] and pre-B/pro-B/common acute lymphoblastic leukemia (ALL). The cells that sustain these leukemias appear to have become directed to generate cells of a particular cell type. An accumulation of erythroid precursors and myeloblasts in AML FAB-M6 reveals a disease origin in a cell with multi-lineage potential[139,140]. However, the partial differentiation of the leukemic blast cells is restricted to certain pathways, and accordingly disease sub-sets are characterized as myeloblast-rich (FAB-M6A), proerythroblast-rich (FAB-M6B) and myeloblast- and proerythroblast-rich (FAB-M6C)[135].

In the case of childhood ALL, including pre-B ALL, pro-B ALL, and common ALL (c-ALL), there are arguments to support a disease origin in either a cell that is committed to B lymphocyte development[141,142] or a cell that is more stem cell like[143]. The former notion is a long-held viewpoint. In favor of the latter cellular origin is that c-ALL-derived cells lacking the B-lineage markers CD10 and CD19 and expressing the stem cell marker CD34 can give rise to c-ALL and pre-B ALL when transplanted into mice[143]. The argument about the precise "target" cell that is transformed in childhood ALL could be set aside as versatility of lineage options extends to progenitor cells, as illustrated by the cytokine-mediated redirection of CLPs to the myeloid lineage[49]. Strikingly, the blast cells that accumulate in the blood and bone marrow in childhood ALL are restricted to B-lineage development.

Conclusions and outlook

There is good evidence to support the notions of developmental plasticity in the hematopoietic system and that growth factors can instruct lineage potentials. This leads to interesting questions as to how HSCs establish and, as required for steady-state hematopoiesis, maintain pluripotency, and how a network of factors might participate to control HSC identity and commitment ability to drive contributions to homeostasis and adaption to inflammatory conditions. Presumably to maintain HSC pluripotency, there is the requirement to control genomic stability, in part via controlling intrinsic DNA repair machinery. As mentioned above vWF^{+} HSCs require Tpo for their maintenance. This growth factor also plays a role in regulating HSC genomic stability, since the efficiency of DNA-Protein Kinase-dependent DNA repair, in response to DNA damage, is increased by Tpo[144]. Tpo-induced activation of ERK and NF-κB in HSCs is important to damage repair[145]. As to Tpo playing a role to ensure chromosomal integrity, the plot thickens regarding the roles of growth factors.

Ineffective DNA repair of genomic instability in LIC and/or LSC might curtail the availability of lineage options of these cells. Global genomic instability is unusual for AML[146], and for both AML and pro-B ALL, a very small number of mutations are required to generate the leukemia[147]. Genomic instability is a feature of the chronic phase of chronic myeloid leukemia, resulting in BCR-ABL-1 mutations that encode resistance to the tyrosine kinase inhibitors (for example imatinib) used to treat the disease. Importantly, instability has been postulated to occur in primitive leukemia progenitor cells in patients who have not been treated with tyrosine kinase inhibitors. Also, it has been suggested that the instability is due to high levels of DNA damage by reactive oxygen species and inefficient/unfaithful repair of DNA double-strand breaks that lead to chromosomal aberrations[148].

In summary, for normal stem and progenitor cells, the versatile nature of the pair-wise model provides an appropriate template for the persuasive action of growth factors to shape hematopoiesis. As yet, we do not know whether LIC and/or LSC are as or less versatile in terms of lineage options than their normal counterparts. Selective instability in genomic elements encoding controls on lineage decision-making (for example transcription factors and signaling molecules) might restrict the versatility of LIC/LSC.

Declaration of interest

The authors alone are responsible for the content and writing of the paper.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013 under Research Executive Agency grant agreement № 315902. C.J.M., L.A.S. and L.vM. gratefully acknowledge receipt of a Marie Curie Research Associate post. G.B., K.T., R.C. and A.R. are partners within the Marie Curie Initial Training Network DECIDE (Decision-making within cells and differentiation entity therapies).

References

1. Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* 1997; 91:661–72.
2. Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 2000;404:193–7.

3. Weissman IL, Anderson DJ, Gage F. Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. *Annu Rev Cell Dev Biol* 2001;17:387–403.
4. Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of mouse hematopoietic stem cells. *Science* 1988;241:58–62.
5. Osawa M, Hanada K, Hamada H, Nakauchi H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* 1996;273:242–5.
6. Spangrude GJ, Brooks DM, Tumas DB. Long-term repopulation of irradiated mice with limiting numbers of purified hematopoietic stem cells: in vivo expansion of stem cell phenotype but not function. *Blood* 1995;85:1006–16.
7. Matsuoka S, Ebihara Y, Xu M, et al. CD34 expression on long-term repopulating hematopoietic stem cells changes during developmental stages. *Blood* 2001;97:419–25.
8. Christensen JL, Weissman IL. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci USA* 2001;98:14541–6.
9. Yang L, Bryder D, Adolfsson J, et al. Identification of Lin(-)Sca1(+)kit(+)CD34(+)Flt3-short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* 2005;105:2717–23.
10. Oguro H, Ding L, Morrison SJ. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* 2013;13:102–16.
11. Ema H, Morita Y, Suda T. Heterogeneity and hierarchy of hematopoietic stem cells. *Exp Hematol* 2014;42:74–82.
12. Challen GA, Boles NC, Chambers SM, Goodell MA. Distinct hematopoietic stem cell subtypes are differentially regulated by TGF-beta1. *Cell Stem Cell* 2010;6:265–78.
13. Morita Y, Ema H, Nakauchi H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med* 2010;207:1173–82.
14. Gekas C, Graf T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood* 2013;121:4463–72.
15. Shimazu T, Iida R, Zhang Q, et al. CD86 is expressed on murine hematopoietic stem cells and denotes lymphopoietic potential. *Blood* 2012;119:4889–97.
16. Beerman I, Bhattacharya D, Zandi S, et al. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proc Natl Acad Sci USA* 2010;107:5465–70.
17. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* 2013;502:232–6.
18. Cumano A, Paige CJ, Iscove NN, Brady G. Bipotential precursors of B cells and macrophages in murine fetal liver. *Nature* 1992;356:612–15.
19. Montecino-Rodriguez E, Leathers H, Dorshkind K. Bipotential B-macrophage progenitors are present in adult bone marrow. *Nat Immunol* 2001;2:83–8.
20. Balciunaite G, Ceredig R, Massa S, Rolink AG. A B220+ CD117+ CD19-hematopoietic progenitor with potent lymphoid and myeloid developmental potential. *Eur J Immunol* 2005;35:2019–30.
21. Adolfsson J, Mansson R, Buza-Vidas N, et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* 2005;121:295–306.
22. Ceredig R, Rolink AG, Brown G. Models of haematopoiesis: seeing the wood for the trees. *Nat Rev Immunol* 2009;9:293–300.
23. Brown G, Ceredig R. Lineage determination in haematopoiesis: quo vadis? *Trends Immunol* 2009;30:465–6.
24. Brown G, Hughes PJ, Michell RH, et al. The sequential determination model of hematopoiesis. *Trends Immunol* 2007;28:442–8.
25. Brown G, Hughes PJ, Michell RH, Ceredig R. The versatility of haematopoietic stem cells: implications for leukaemia. *Crit Rev Clin Lab Sci* 2010;47:171–80.
26. Ishikawa F, Niiro H, Iino T, et al. The developmental program of human dendritic cells is operated independently of conventional myeloid and lymphoid pathways. *Blood* 2007;110:3591–660.
27. Manz MG, Traver D, Miyamoto T, et al. Dendritic cell potentials of early lymphoid and myeloid progenitors. *Blood* 2001;97:3333–41.
28. Gauvreau GM, Ellis AK, Denburg JA. Haemopoietic processes in allergic disease: eosinophil/basophil development. *Clin Exp Allergy* 2009;39:1297–306.
29. Pang L, Weiss MJ, Poncz M. Megakaryocyte biology and related disorders. *J Clin Invest* 2005;115:3332–8.
30. Fogg DK, Sibon C, Miled C, et al. A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* 2006;311:83–7.
31. Nozad Charoudeh H, Tang Y, Cheng M, et al. Identification of an NK/T cell-restricted progenitor in adult bone marrow contributing to bone marrow-and thymic-dependent NK cells. *Blood* 2010;116:183–92.
32. Luc S, Buza-Vidas N, Jacobsen SE. Delineating the cellular pathways of hematopoietic lineage commitment. *Semin Immunol* 2008;20:213–20.
33. Katsura Y. Redefinition of lymphoid progenitors. *Nat Rev Immunol* 2002;2:127–32.
34. Ye M, Graf T. Early decisions in lymphoid development. *Curr Opin Immunol* 2007;19:123–8.
35. Sun J, Ramos A, Chapman B, et al. Clonal dynamics of native haematopoiesis. *Nature* 2014;514:322–7.
36. Guimaraes JE, Francis GE, Bol SJ, et al. Differentiation restriction in the neutrophil-granulocyte, macrophage, eosinophil-granulocyte pathway: analysis by equilibrium density centrifugation. *Leuk Res* 1982;6:791–800.
37. Braunstein M, Rajkumar P, Claus CL, et al. HEBAlt enhances the T-cell potential of fetal myeloid-biased precursors. *Int Immunol* 2010;22:963–72.
38. Shankland M. Differentiation of the O and P cell lines in the embryo of the leech. II. Genealogical relationship of descendant pattern elements in alternative developmental pathways. *Dev Biol* 1987;123:97–107.
39. Porritt HE, Rumfelt LL, Tabrizifard S, et al. Heterogeneity among DN1 prothymocytes reveals multiple progenitors with different capacities to generate T cell and non-T cell lineages. *Immunity* 2004;20:735–45.
40. Benz C, Bleul CC. A multipotent precursor in the thymus maps to the branching point of the T versus B lineage decision. *J Exp Med* 2005;202:21–31.
41. Balciunaite G, Ceredig R, Rolink AG. The earliest subpopulation of mouse thymocytes contains potent T, significant macrophage, and natural killer cell but no B-lymphocyte potential. *Blood* 2005;105:1930–6.
42. Taghon T, Yui MA, Rothenberg EV. Mast cell lineage diversion of T lineage precursors by the essential T lineage transcription factor GATA-3. *Nat Immunol* 2007;8:845–55.
43. Nakahata T, Ogawa M. Clonal origin of murine hemopoietic colonies with apparent restriction to granuclocyte-macrophage-megakaryocyte (GMM) differentiation. *J Cell Physiol* 1982;111:239–46.
44. Ogawa M, Porter PN, Nakahata T. Renewal and commitment to differentiation of hemopoietic stem cells (an interpretive review). *Blood* 1983;61:823–9.
45. Suda T, Suda J, Ogawa M. Disparate differentiation in mouse hemopoietic colonies derived from paired progenitors. *Proc Natl Acad Sci USA* 1984;81:2520–4.
46. Enver T, Heyworth CM, Dexter TM. Do stem cells play dice? *Blood* 1998;92:348–51; discussion 52.
47. Metcalf D. Hematopoietic cytokines. *Blood* 2008;111:485–91.
48. Robb L. Cytokine receptors and hematopoietic differentiation. *Oncogene* 2007;26:6715–23.
49. Kondo M, Scherer DC, Miyamoto T, et al. Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. *Nature* 2000;407:383–6.
50. Rieger MA, Hoppe PS, Smejkal BM, et al. Hematopoietic cytokines can instruct lineage choice. *Science* 2009;325:217–18.
51. Metcalf D, Burgess AW. Clonal analysis of progenitor cell commitment of granulocyte or macrophage production. *J Cell Physiol* 1982;111:275–83.
52. Metcalf D. Lineage commitment of hemopoietic progenitor cells in developing blast cell colonies: influence of colony-stimulating factors. *Proc Natl Acad Sci USA* 1991;88:11310–14.

53. Mossadegh-Keller N, Sarrazin S, Kandalla PK, et al. M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature* 2013;497:239–43.
54. Grover A, Mancini E, Moore S, et al. Erythropoietin guides multipotent hematopoietic progenitor cells toward an erythroid fate. *J Exp Med* 2014;211:181–8.
55. Lyman SD, James L, Vanden Bos T, et al. Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase receptor: a proliferative factor for primitive hematopoietic cells. *Cell* 1993;75: 1157–67.
56. Matthews W, Jordan CT, Wiegand GW, et al. A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations. *Cell* 1991;65:1143–52.
57. Weisberg E, Barrett R, Liu Q, et al. FLT3 inhibition and mechanisms of drug resistance in mutant FLT3-positive AML. *Drug Resist Updat* 2009;12:81–9.
58. Wander SA, Levis MJ, Fathi AT. The evolving role of FLT3 inhibitors in acute myeloid leukemia: quizartinib and beyond. *Ther Adv Hematol* 2014;5:65–77.
59. Rasko JE, Metcalf D, Rossner MT, et al. The flt3/flk-2 ligand: receptor distribution and action on murine haemopoietic cell survival and proliferation. *Leukemia* 1995;9:2058–66.
60. Adolfsson J, Borge OJ, Bryder D, et al. Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* 2001;15:659–69.
61. Karsunky H, Merad M, Cozzio A, et al. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells in vivo. *J Exp Med* 2003;198:305–13.
62. Sitnicka E, Brakebusch C, Martensson IL, et al. Complementary signaling through flt3 and interleukin-7 receptor alpha is indispensable for fetal and adult B cell genesis. *J Exp Med* 2003;198: 1495–506.
63. Ahsberg J, Tsapogas P, Qian H, et al. Interleukin-7-induced Stat-5 acts in synergy with Flt-3 signaling to stimulate expansion of hematopoietic progenitor cells. *J Biol Chem* 2010;285: 36275–84.
64. Balciunaite G, Ceredig R, Fehling HJ, et al. The role of Notch and IL-7 signaling in early thymocyte proliferation and differentiation. *Eur J Immunol* 2005;35:1292–300.
65. Veiby OP, Lyman SD, Jacobsen SE. Combined signaling through interleukin-7 receptors and flt3 but not c-kit potently and selectively promotes B-cell commitment and differentiation from uncommitted murine bone marrow progenitor cells. *Blood* 1996; 88:1256–65.
66. Mackarehtschian K, Hardin JD, Moore KA, et al. Targeted disruption of the flk2/flt3 gene leads to deficiencies in primitive hematopoietic progenitors. *Immunity* 1995;3:147–61.
67. McKenna HJ, Stocking KL, Miller RE, et al. Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* 2000;95:3489–97.
68. Tsapogas P, Swee LK, Nusser A, et al. *In vivo* evidence for an instructive role of fms-like tyrosine kinase-3 (FLT3) ligand in hematopoietic development. *Haematologica* 2014;99:638–46.
69. Ceredig R, Rauch M, Balciunaite G, Rolink AG. Increasing Flt3L availability alters composition of a novel bone marrow lymphoid progenitor compartment. *Blood* 2006;108:1216–22.
70. Ceredig R, Rolink AG. The key role of IL-7 in lymphopoiesis. *Semin Immunol* 2012;24:159–64.
71. Corfe SA, Paige CJ. The many roles of IL-7 in B cell development; mediator of survival, proliferation and differentiation. *Semin Immunol* 2012;24:198–208.
72. Gonzalez-Garcia S, Garcia-Peydro M, Alcain J, Toribio ML. Notch1 and IL-7 receptor signalling in early T-cell development and leukaemia. *Curr Top Microbiol Immunol* 2012;360: 47–73.
73. Namen AE, Schmierer AE, March CJ, et al. B cell precursor growth-promoting activity. Purification and characterization of a growth factor active on lymphocyte precursors. *J Exp Med* 1988; 167:988–1002.
74. Peschon JJ, Morrissey PJ, Grabstein KH, et al. Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J Exp Med* 1994;180:1955–60.
75. Puel A, Ziegler SF, Buckley RH, Leonard WJ. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. *Nat Genet* 1998;20:394–7.
76. von Freeden-Jeffry U, Vieira P, Lucian LA, et al. Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J Exp Med* 1995;181:1519–26.
77. Parrish YK, Baez I, Milford TA, et al. IL-7 dependence in human B lymphopoiesis increases during progression of ontogeny from cord blood to bone marrow. *J Immunol* 2009;182:4255–66.
78. Tsapogas P, Zandi S, Ahsberg J, et al. IL-7 mediates Ebf-1-dependent lineage restriction in early lymphoid progenitors. *Blood* 2011;118:1283–90.
79. Miller JP, Izon D, DeMuth W, et al. The earliest step in B lineage differentiation from common lymphoid progenitors is critically dependent upon interleukin 7. *J Exp Med* 2002;196:705–11.
80. Kikuchi K, Lai AY, Hsu CL, Kondo M. IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *J Exp Med* 2005;201:1197–203.
81. Hirokawa S, Sato H, Kato I, Kudo A. EBF-regulating Pax5 transcription is enhanced by STAT5 in the early stage of B cells. *Eur J Immunol* 2003;33:1824–9.
82. O'Riordan M, Grosschedl R. Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. *Immunity* 1999;11:21–31.
83. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 1999;401:556–62.
84. Rolink AG, Nutt SL, Melchers F, Busslinger M. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature* 1999;401:603–6.
85. Rolink A, Kudo A, Melchers F. Long-term proliferating early pre-B-cell lines and clones with the potential to develop to surface immunoglobulin-positive, mitogen-reactive B-cells in vitro and in vivo. *Biochem Soc Trans* 1991;19:275–6.
86. Malin S, McManus S, Cobaleda C, et al. Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. *Nat Immunol* 2010;11: 171–9.
87. Milne CD, Fleming HE, Paige CJ. IL-7 does not prevent pro-B/pre-B cell maturation to the immature/sIgM(+) stage. *Eur J Immunol* 2004;34:2647–55.
88. Milne CD, Paige CJ. IL-7: a key regulator of B lymphopoiesis. *Semin Immunol* 2006;18:20–30.
89. Bertolino E, Reddy K, Medina KL, et al. Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. *Nat Immunol* 2005;6:836–43.
90. Chowdhury D, Sen R. Stepwise activation of the immunoglobulin mu heavy chain gene locus. *EMBO J* 2001;20:6394–403.
91. Corcoran AE, Riddell A, Krooshoop D, Venkitaraman AR. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* 1998;391:904–7.
92. Mandal M, Powers SE, Maienschein-Cline M, et al. Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. *Nat Immunol* 2011;12:1212–20.
93. Rolink AG, Winkler T, Melchers F, Andersson J. Precursor B cell receptor-dependent B cell proliferation and differentiation does not require the bone marrow or fetal liver environment. *J Exp Med* 2000;191:23–32.
94. Grawunder U, Haasner D, Melchers F, Rolink A. Rearrangement and expression of kappa light chain genes can occur without mu heavy chain expression during differentiation of pre-B cells. *Int Immunol* 1993;5:1609–18.
95. Johnson K, Hashimshony T, Sawai CM, et al. Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. *Immunity* 2008;28:335–45.
96. Marshall AJ, Fleming HE, Wu GE, Paige CJ. Modulation of the IL-7 dose-response threshold during pro-B cell differentiation is dependent on pre-B cell receptor expression. *J Immunol* 1998;161: 6038–45.
97. Clark MR, Mandal M, Ochiai K, Singh H. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat Rev Immunol* 2014;14:69–80.

98. Schwarz BA, Sambandam A, Maillard I, et al. Selective thymus settling regulated by cytokine and chemokine receptors. *J Immunol* 2007;178:2008–17.
99. Schlenner SM, Madan V, Busch K, et al. Fate mapping reveals separate origins of T cells and myeloid lineages in the thymus. *Immunity* 2010;32:426–36.
100. Ceredig R. Fates and potentials of thymus-seeding progenitors. *Nat Immunol* 2012;13:309–10.
101. Gonzalez-Garcia S, Garcia-Peydro M, Martin-Gayo E, et al. CSL-MAML-dependent Notch1 signaling controls T lineage-specific IL-7R{alpha} gene expression in early human thymopoiesis and leukemia. *J Exp Med* 2009;206:779–91.
102. Petrie HT, Zuniga-Pflucker JC. Zoned out: functional mapping of stromal signaling microenvironments in the thymus. *Annu Rev Immunol* 2007;25:649–79.
103. Radtke F, Wilson A, Stark G, et al. Deficient T cell fate specification in mice with an induced inactivation of Notch1. *Immunity* 1999;10:547–58.
104. Takahama Y. Journey through the thymus: stromal guides for T-cell development and selection. *Nat Rev Immunol* 2006;6: 127–35.
105. Garcia-Peydro M, de Yebenes VG, Toribio ML. Notch1 and IL-7 receptor interplay maintains proliferation of human thymic progenitors while suppressing non-T cell fates. *J Immunol* 2006; 177:3711–20.
106. Plum J, De Smedt M, Leclercq G, et al. Interleukin-7 is a critical growth factor in early human T-cell development. *Blood* 1996;88: 4239–45.
107. Shortman K, Egerton M, Spangrude GJ, Scollay R. The generation and fate of thymocytes. *Semin Immunol* 1990;2:3–12.
108. De Smedt M, Reynvoet K, Kerre T, et al. Active form of Notch imposes T cell fate in human progenitor cells. *J Immunol* 2002; 169:3021–9.
109. Sudo T, Nishikawa S, Ohno N, et al. Expression and function of the interleukin 7 receptor in murine lymphocytes. *Proc Natl Acad Sci USA* 1993;90:9125–9.
110. Maraskovsky E, O'Reilly LA, Teepe M, et al. Bcl-2 can rescue T lymphocyte development in interleukin-7 receptor-deficient mice but not in mutant rag-1 / mice. *Cell* 1997;89:1011–19.
111. Kang J, Volkmann A, Raulet DH. Evidence that gammadelta versus alphabeta T cell fate determination is initiated independently of T cell receptor signaling. *J Exp Med* 2001;193:689–98.
112. Durum SK, Candeias S, Nakajima H, et al. Interleukin 7 receptor control of T cell receptor gamma gene rearrangement: role of receptor-associated chains and locus accessibility. *J Exp Med* 1998;188:2233–41.
113. Kondo M, Akashi K, Domen J, et al. Bcl-2 rescues T lymphopoiesis, but not B or NK cell development, in common gamma chain-deficient mice. *Immunity* 1997;7:155–62.
114. Ikawa T, Hirose S, Masuda K, et al. An essential developmental checkpoint for production of the T cell lineage. *Science* 2010;329: 93–6.
115. Yu Q, Park JH, Doan LL, et al. Cytokine signal transduction is suppressed in preselection double-positive thymocytes and restored by positive selection. *J Exp Med* 2006;203:165–75.
116. Munitic I, Williams JA, Yang Y, et al. Dynamic regulation of IL-7 receptor expression is required for normal thymopoiesis. *Blood* 2004;104:4165–72.
117. Tussiwand R, Engdahl C, Gehre N, et al. The preTCR-dependent DN3 to DP transition requires Notch signaling, is improved by CXCL12 signaling and is inhibited by IL-7 signaling. *Eur J Immunol* 2011;41:3371–80.
118. Gascoigne NR. CD8+ thymocyte differentiation: T cell two-step. *Nat Immunol* 2010;11:189–90.
119. Park JH, Adoro S, Guinter T, et al. Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. *Nat Immunol* 2010;11:257–64.
120. Schluns KS, Kieper WC, Jameson SC, Lefrancois L. Interleukin-7 mediates the homeostasis of naive and memory CD8 T cells *in vivo*. *Nat Immunol* 2000;1:426–32.
121. Tani-ichi S, Shimba A, Wagatsuma K, et al. Interleukin-7 receptor controls development and maturation of late stages of thymocyte subpopulations. *Proc Natl Acad Sci USA* 2013;110:612–17.
122. Kincade PW, Medina KL, Payne KJ, et al. Early B-lymphocyte precursors and their regulation by sex steroids. *Immunol Rev* 2000; 175:128–37.
123. Bosco N, Ceredig R, Rolink A. Transient decrease in interleukin-7 availability arrests B lymphopoiesis during pregnancy. *Eur J Immunol* 2008;38:381–90.
124. Phuc LH, Papiernik M, Berrih S, Duval D. Thymic involution in pregnant mice. *I. Characterization of the remaining thymocyte subpopulations. Clin Exp Immunol* 1981;44:247–52.
125. Li Y, Esain V, Teng L, et al. Inflammatory signalling regulates embryonic hematopoietic stem and progenitor cell production. *Genes Dev* 2014;28:2597–612.
126. Dick JE. Stem cell concepts renew cancer research. *Blood* 2008; 112:4793–807.
127. Kreso A, Dick JE. Evolution of the cancer stem cell model. *Cell Stem Cell* 2014;14:275–91.
128. Lane SW, Gilliland DG. Leukemia stem cells. *Semin Cancer Biol* 2010;20:71–6.
129. Buzzai M, Licht JD. New molecular concepts and targets in acute myeloid leukemia. *Curr Opin Hematol* 2008;15:82–7.
130. Krause DS, Van Etten RA. Right on target: eradicating leukemic stem cells. *Trends Mol Med* 2007;13:470–81.
131. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 1997;3:730–7.
132. Sanchez-Garcia I, Vicente-Duenas C, Cobaleda C. The theoretical basis of cancer-stem-cell-based therapeutics of cancer: can it be put into practice? *Bioessays* 2007;29:1269–80.
133. Greaves M. Darwin and evolutionary tales in leukemia. The Ham-Wasserman Lecture. *Hematology Am Soc Hematol Educ Program.* 2009;1:3–12.
134. Greaves M. Cancer stem cells: back to Darwin? *Semin Cancer Biol* 2010;20:65–70.
135. Brown G, Hogg N, Greaves M. Candidate leukaemia-specific antigen in man. *Nature* 1975;258:454–6.
136. Greaves MF, Brown G, Capellaro D, et al. Immunological approaches to the identification of leukaemia cells. In: Wybran J, Staquet MJ, eds. *Clinical Tumour Immunology*. Oxford: Pergamon Press, 1976:115–22.
137. Brown G, Hughes PJ, Ceredig R. The versatile landscape of haematopoiesis: are leukaemia stem cells as versatile? *Crit Rev Clin Lab Sci* 2012;49:232–40.
138. Brown G, Hughes PJ, Ceredig R, Michell RH. Versatility and nuances of the architecture of haematopoiesis – implications for the nature of leukaemia. *Leuk Res* 2012;36:14–22.
139. Davey FR, Abraham Jr N, Brunetto VL, et al. Morphologic characteristics of erythroleukemia (acute myeloid leukemia; FAB-M6): a CALGB study. *Am J Hematol* 1995;49:29–38.
140. Kowal-Vern A, Mazzella FM, Cotelingam JD, et al. Diagnosis and characterization of acute erythroleukemia subsets by determining the percentages of myeloblasts and proerythroblasts in 69 cases. *Am J Hematol* 2000;65:5–13.
141. Greaves MF. Differentiation-linked leukemogenesis in lymphocytes. *Science* 1986;234:697–704.
142. Greaves M. Molecular genetics, natural history and the demise of childhood leukaemia. *Eur J Cancer* 1999;35:1941–53.
143. Cox CV, Blair A. A primitive cell origin for B-cell precursor ALL? *Stem Cell Rev* 2005;1:189–96.
144. de Laval B, Pawlikowska P, Petit-Cocault L, et al. Thrombopoietin-increased DNA-PK-dependent DNA repair limits hematopoietic stem and progenitor cell mutagenesis in response to DNA damage. *Cell Stem Cell* 2013;12:37–48.
145. de Laval B, Pawlikowska P, Barbieri D, et al. Thrombopoietin promotes NHEJ DNA repair in hematopoietic stem cells through specific activation of Erk and NF-kappaB pathways and their target, IEX-1. *Blood* 2014;123:509–19.
146. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell* 2012;150: 264–78.
147. Dobbins SE, Sherborne AL, Ma YP, et al. The silent mutational landscape of infant MLL-AF4 pro-B acute lymphoblastic leukemia. *Genes Chromosom Cancer* 2013;52:954–60.
148. Bolton-Gillespie E, Schemionek M, Klein HU, et al. Genomic instability may originate from imatinib-refractory chronic myeloid leukemia stem cells. *Blood* 2013;121:4175–83.

10. Acknowledgements

I want to thank all people who supported me during my time as a PhD student.

My first thanks go to Ton for allowing me to take part of your brilliant team, it has been a great honour! Thank you for your guidance and sharing your enormous expertise and knowledge as well as for maintaining always with a laugh the good environment characteristic of the lab. Moreover, I want to thank you for encouraging me to explore a new scientific dimension, the bioinformatics and for providing me the tools, time and comprehension to learn and grow in this field. In summary, thanks for your always support!

Secondly I would like to thank Daniela and Rhodri, who kindly agreed to become members of my PhD committee. I received great input and feedback during numerous lab meetings as well as DECIDE meetings. Thank you, Rhodri, for your time in Basel, for our discussions and for your enormously helpful corrections on the thesis.

The third, but not less important, big thank you goes to all current and former members of “Team Ton”. It has been a great team with an enviable atmosphere and a helping hand whenever needed as well as scientific advice. Thank you Fabian, Lilly, Panos, Patrick, Stefan, Marianne, Corinne, Mike, Giusy, Jonas, Hannie, Nicole, Simon, Ilda, Sofia and Jan. Specially, I want to thank Lilly, who has been also part of the DECIDE network, for all the moments and support to each other; Panos, for all the discussions about science and the world in general, for your scientific input, encouraging words and for the time dedicated to fine tune the thesis; Patrick, for all the fun and for your enormous and contagious positivism and enthusiasm in the lab, you have been a great support! Moreover, I also want to thank Daniela’s and Roxanne’s group for sharing the lab seminars and helpful discussions. Besides the lab work I enjoyed all the fun we had at Christmas parties, lab outings, coffee breaks, late beers in the lab, retreats, the use of the hand-made “signs” etc

I would also like to thank the DECIDE network. I feel very proud and fortunate to have taken part in this PhD Marie Curie programme. I really want to thank all the people that have made this network possible because it has been a huge opportunity for us. It was great to bring all the fellows, supervisors and partners of the DECIDE network together once a year in a different country. Thank you to all the fellows, I am so happy to have met you and share our parallel PhD experiences at different countries. Let's keep in touch with each other!

Finally, to all the people outside science who has supported me. My girls, Flavia, Marina and Marian, for your valuable distance friendship, understanding and encouraging me on keeping going during my PhD. Also for our yearly holidays where I could refill the energy and keep the happiness and motivation. Thanks to all the people who have contributed to my happy stay in Basel: Basilenyus, DBM community, Marwa, Lina, Maria etc. but above all, to Fran. Cariño, no tengo palabras para expresarte el agradecimiento que siento hacia ti. Sin lugar a duda has sido la persona que has vivido más de cerca el inicio, desarrollo y culminación de mi tesis doctoral. Tu apoyo ha sido fundamental, siempre has tenido palabras de ánimo y comprensión (así como hechos). Vine a Basel para embarcarme en el proyecto de hacer un doctorado, pero gracias a ti esta experiencia se ha convertido en algo inolvidable, juntos hemos descubierto un nuevo país, nos hemos adaptado y hemos tirado hacia delante. Eres esa persona esencial que necesito en mi vida para futuros proyectos y a la que también pienso apoyar en todos los que tengas. Gracias de corazón por ser ante todo mi amigo y confidente, el mejor compañero de viaje. Te quiero.

També vull agrair a tots els amics, sobretot de Mallorca, per entendre lo enfeinada que he estat escrivint. No vull acabar sense agraïr des de lo més profund del cor a tota la meva familia i, sobretot en es meus pares i germà. Papas moltes gracis per haver confiat sempre amb jo i per haver facilitat sempre que compleixi es meus somnis. No sería on sóc si no fos per voltros. Per molt lluny que estiguem, sempre vos he sentit a prop. A més, aquests darrers dies escrivint des de casa (Mallorca) m'han donat energies extres per acabar sa tesi. Ha estat molt d'esforç però ha valgut la pena. Moltes gracis per tot, vos estim.